STUDIES OF THE PHYSICAL STABILITY OF GLP-1 & CHEMICALLY MODIFIED FORMS OF GLP-1

Frederik Johannes Becher

Peterhouse

The Jackson Laboratory

Department of Chemistry

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

November 2020

This thesis is dedicated to my grandparents

“Ich habe keine besondere Begabung, sondern bin nur leidenschaftlich neugierig.”

Albert Einstein

“Ideen, wie absolute Gewißheit, absolute Genauigkeit, endgültige Wahrheit und so fort, sind Erfindungen der Einbildungskraft und haben in der Wissenschaft nichts zu suchen.”

Max Born

DECLARATION

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university of institution for any degree, diploma, or other qualification.

In accordance with the Department of Chemistry guidelines, this thesis is does not exceed 60,000 words.

Signed:

Date: 02/11/2020

Frederik Johannes Becher

Cambridge

SUMMARY

Biopharmaceuticals, including proteins and peptides, are becoming increasingly important as therapeutic agents. However, the clinical use of protein and peptide therapeutics is still restricted due to undesirable properties such as their ability to self-assemble and aggregate. These properties not only influence production processes and storage but can also have adverse immunogenic effects within humans. To enhance their use, it is of great importance to understand and be able to manipulate their aggregation behaviour. *In vivo,* protein- and peptide-based drugs suffer from different problems, such as short lifetimes and low stability. Lipidation of peptides is widely established as a means of increasing stability *in vivo*. However, relatively little is known about the effect of lipidation on peptide self-assembly and other aggregation phenomena *in vitro*.

The aim of this study is to develop a better understanding of the mechanism of aggregation and amyloid fibrillation of the therapeutic peptide GLP-1 and chemically modified forms of GLP-1. GLP-1, a metabolic hormone, has the ability to decrease blood sugar levels in a glucose-dependent manner by enhancing the secretion of insulin. Previous measurements of GLP-1 aggregation revealed that, at certain pH values, unusual behaviour is observed that has established that the standard nucleation-polymerization mechanism is insufficient to fully describe the reaction under these conditions.

To study the aggregation mechanism of GLP-1 and chemically modified forms of the peptide, aggregation kinetics were measured over a wide range of different conditions. In addition, many biophysical techniques, such as AFM, SEM, far-UV CD, FT-IR, λ_{max}, ANS, *ex situ* ThT and DLS, were employed to probe the structure, size and properties of species in solution during aggregation.

The influence of amidation at the C-terminus of GLP-1 on physical stability was assessed and significant differences to GLP-1 were found. The study of this relatively small modification generates a better understanding of the interdependence of net charge, solubility and secondary structure on the aggregation kinetics. The results of these studies also provide further evidence that peptides belonging to the GLP-1 family can form off-pathway oligomeric species that have a significant impact on the aggregation kinetics. Two lipidated analogues of C-terminally amidated GLP-1 (Am-GLP-1) were also studied and the results analysed in detail and compared to those obtained for GLP-1 and Am-GLP-1. Both lipidated peptides show a strong, nearly switch-like, pH dependence. Surprisingly, it was also shown that the amidation of the C-terminus had a bigger influence on the secondary structure of the peptides in comparison to the lipidation.

ACKNOWLEDGEMENTS

During my PhD, I had the support of a broad range of people – combinations of my family, friends, supervisors and co-workers all deserve my gratitude.

Starting with my two supervisors, Prof Sophie Jackson and Dr Ana dos Santos, who helped me discover the exciting world of peptide aggregation. And it's been an honour to work with you and do so at such magnificent places – the Chemistry Department at the University of Cambridge and the laboratory of MedImmune (now AstraZeneca - AZ).

I want to say a big thank you for all the support, positive challenges and scientific discussions. I have learned and will continue to learn a lot from both of you.

I also want to acknowledge the entire Jackson Group. Special thanks go to Karolina for laying the foundation to my project. Your continued help, always introducing me to different techniques and experiments, was priceless. I appreciated all questions, discussions of new data and lunch-time chats with Eva, Katie, Jack and Carolina.

During my time at AZ, I have also had the support of the whole Formulation group (later DFDD group). Special thanks go to Shahid Udin and Steven Bishop for their support and interest in my work. Like many members of the Jackson Group, I was fortunate enough to receive help of many AZ employees. Specifically, I want to highlight Anne-Laure for her support with FT-IR and Elise for her support with LC-MS. Also, special thanks to Sophie for answering my many questions in the laboratory. In general, every person I met at AZ, without any exception, was absolutely helpful and supportive in my studies. For this, I am very lucky and truly grateful. And finally, all the other PhD students at AZ deserve an honourable mention. It was great to share my time with both Maria-Laura and Marcello – there was nothing we couldn't solve together in the lab.

In this context, I am greatly appreciative of the NanoScience group of Prof Sir Mark Welland which welcomed me during my AFM measurements. Special thanks go to Simon for explaining to me in detail everything I had to know about the instrument. I also thank Sonja and Géraldine, who were also always

there if I had any questions. Altogether, we had a great time in Washington and Gaithersburg at MedImmune.

Finally, I want to mention and thank George for answering all my questions regarding NEF-fitting.

There is something in Cambridge which makes your life as a PhD student a bit more special than usual and that is the lovely college life. I am grateful that I was in the fortunate situation to become a member of Peterhouse, the oldest college of Cambridge. Great thanks go therefore to the entire college staff which I got to know very well during my year in the MCR committee. They all always made us feel like we were at home!

I also want to thank to the whole MCR committee during my time as social secretary. It is not exaggerated to say that we all enriched the social life of Peterhouse as much as we could! Many thanks to Leo, Laura, Simone, Monique, Frederik and Jamie. I am sure one day Medusa will join us again to a fantastic formal hall. An essential part of my time in Cambridge was enjoying the fascinating conversations and beautiful walks along the River Cam with William and Simon. I really appreciated your thoughts!

Another committee that played a major part of my life in Cambridge was the committee of the German Society. I thank all of the co-members for some of the finest memories you can imagine.

I must also thank my best friends, Alexander, Arni and Philipp. It was a pleasure to show you around in Cambridge. I am really looking forward to our next adventures!

It was something extraordinary to have Christian and Maija, two of my closest friends, around London and Cambridge. Both played an essential part in making me feel at home in the UK.

In the end, I was lucky to have a really close friend offer me shelter during Covid-19 pandemic and at one of the most inspiring places one can imagine. I am really grateful Ann for providing me with a room at your little horse range. And many thanks to Marc, Leana, Sven and Saskia – they welcomed me warmly and made the time memorable despite all the work writing.

Last but not least, I want to say thank you to my family. An enormous thank

you goes to my parents, their partners and my brothers with their partners as well. They always support me in every step I take.

Finally, I say thank you to AstraZeneca (MedImmune) and Peterhouse for financing the project.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ALL FIGURES TAKEN FROM OTHER SOURCES ARE DONE SO WITH PERMISSION FROM THE RELEVANT JOURNALS.

List of Abbreviations and Acronyms

LIST OF APPENDICES

1 INTRODUCTION

In recent years, biopharmaceuticals specifically peptide therapeutics, have become increasingly important and are now responsible for a considerable share of the pharmaceutic market.[1,2] However, despite many advantages over other types of therapeutic agents such as high bioactivity and specificity, peptides also pose considerable challenges.[3] One of these is associated with their physical stability as many peptides have a tendency to aggregate and precipitate. This has a major impact on their usability and safety: predictable shelf lives are crucial, and it is known that peptide aggregates formed in drug solutions can induce adverse immune responses and lead to a loss of safety and efficacy.[4]

One example of physical instability is polypeptide chain self-assembly into highly regular structures such as amyloid fibrils. The formation of amyloid fibrils is one of the most common processes influencing almost all phases of biological drug development. The importance of understanding these processes becomes clear by the broad range of different protein pharmaceuticals known to form different types of aggregates as shown in Table 1.1.

peptide or protein	length (residues)	conditions	type of aggregates
GLP-1 (7-37)	31	pH 7.5-8.2	fibril[5]
GLP-1 (7-36)-NH_2	30	pH 7.4	fibril[6]
glucagon	29	pH 2.0	fibril[7–9]
glucagon	29	pH 8.5-9.7	fibril[10,11]
salmon calcitonin	32	neutral pH	fibril[12–15]
amylin	37	under different conditions	fibril[16–20]
insulin	51 (31 chain A, 20 chain B)	pH 2.0	fibril[21–23]
		pH 1.1-3.2	fibril[24–27]
		pH 1.5-7.5	fibril[28]
Insulin	51 (31 chain A, 20 chain B)	pH 1.75-2.0	
		conc. < 5 mg ml^{-1}	spherulite[29]
		conc. > 5 mg ml^{-1}	fibril[29]
insulin analogues	51 (31 chain A, 20 chain B)	high temperature and mechanical agitation	precipitation[30]
human parathyroid hormone	84	–	fibril[31]
human interferon a2b		chemical degradation-oxidation	covalent aggregate[32,33]
PYY (3-36) & lipidated derivatives	34	pH 2, 4 & 8	micelles / nanotapes / fibrillar hydrogels[34]
Proleukin		+SDS	formation of micellar-like structures containing approx. 27 molecules[35]
Met-G-CSF		pH 6.9, 37 °C	precipitation[36]

Table 1.1. Some examples of peptide and protein pharmaceuticals known to form different types of aggregates.

To learn more about the formation of amyloid fibrils and the mechanisms by which peptide and protein aggregation occurs, as well as the factors that influence the reaction, numerous and extensive research programmes have been undertaken. This has led to significant progress, especially in the field of peptides and proteins where aggregation *in vivo* is associated with disease states, particularly neurodegenerative diseases.[37] Typical examples of neurodegenerative diseases associated with amyloid aggregates are Alzheimer's Disease (Amyloid-β)[38], Parkinson's disease (α-synuclein)[39,40] and Huntington's disease (Huntington fragments)[41]. A considerable amount of the research on aggregation and self-assembly of disease-related peptides and proteins can be applied to the aggregation of peptide-based therapeutic agents currently developed by the pharmaceutical industry. In general, however, the aggregation pathways of peptides and polypeptide-based drugs coming through the drug pipelines have not been characterised in any depth.

The topic of this Thesis is the study of the physical stability of glucagon-like peptide 1 (GLP-1) and chemically modified forms of GLP-1. To understand the aggregation mechanism and the physical stability of peptide and protein pharmaceuticals, it is essential to understand the influence of mutations and other types of chemical modifications. In this Chapter, a short introduction is provided on the physical stability and aggregation of peptides, kinetics of peptide aggregation, different factors influencing peptide aggregation, the lipidation of peptides and more details about the peptide used throughout this study, GLP-1 and its lipidated analogues.

1.1 A Brief History of Amyloids

The research on amyloids and amyloid fibrils has a long and eventful history. Most likely, the first observation was by Nicolao Fontano, who described an autopsy of a young man with an abscess in his liver and a large spleen filled with white stone.[42] In 1854, Rudolf Virchow described the structures as amyloids because he observed the peculiar reaction of the corpora amylacea of the nervous system with iodine.[43] He adopted the term from Matthias Schleiden, a botanist, who coined the term less than two decades earlier to describe a normal amylaceous constituent of plants.[42] The iodine reaction for

starch was already reported by J.J. Colin and H.F. Gaultier de Clauburg, and independently Stromeyer in 1814.[44] The term amyloid replaced with time the originally more common terms ‘lardaceous’ or ‘waxy’ changes.

In 1859, Carl Friedreich and August Kekulé proved that the waxy spleen they observed consisted neither of amylon nor cellulose but from albumoid compounds based on the high nitrogen content.[42] Why the term amyloidosis still prevailed is not clear, but it is assumed that Virchows standing as a pathologist and the common use of iodine to stain the samples, had an influence.

Aniline based dyes used to recognise amyloid were first described in parallel by André-Victor Cornil, Richard Heschl and Rudolph Jürgen in 1875.[42] These initially used metachromatic stains were eventually replaced by another aniline based stain, Congo Red. Generally used for staining textiles, Congo Red was found to bind avidly to amyloid by Bennhold in 1922.[45]

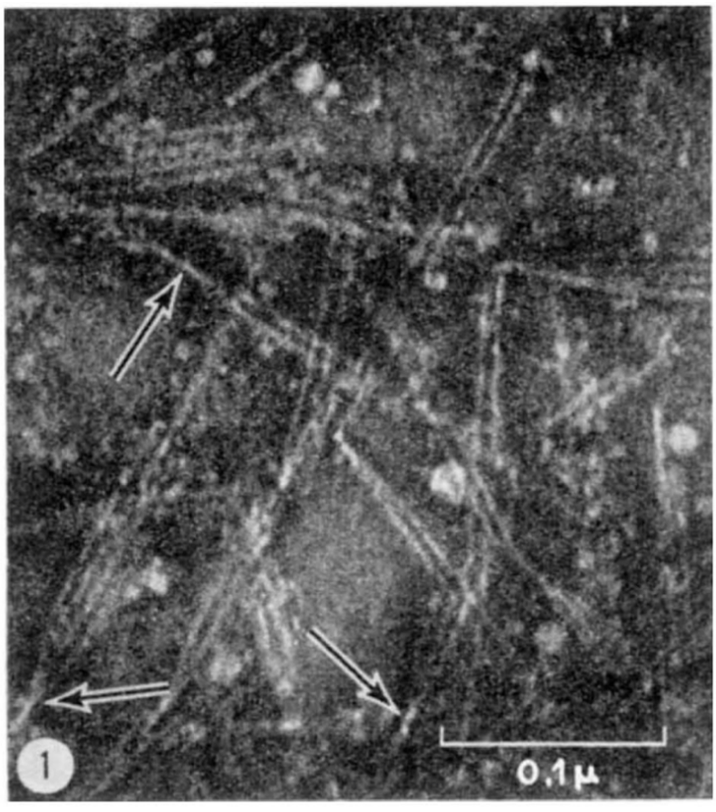


Figure 1.1. Negatively stained amyloid fibrils.

Amyloid fibrils imaged using an electron microscope by Shirahama & Cohen in 1965.[46] Note various combinations and coplanar disposition of filaments and loose random twisting of fibrils (x 250,000).

In the 20th Century, more and more details about the structure of amyloids were discovered. William Astbury took the first diffraction images of amyloid and discovered a fibrous state for many proteins, for example keratin in 1931-33 and pepsin in 1934.[47–50] With the help of electron microscopy, Cohen and Calkins first recognised in 1959 that amyloid demonstrates a non-branching fibrillar structure with a width of 6-13 nm.[51] The first twisting protofilaments were reported only six years later by Shirahana and Cohen,[52] Figure 1.1.

In 1967, Geddes, Parker, Atkins and Beighton also published a paper on X-ray diffraction patterns of naturally occurring cross-β fibrous proteins.[53] Just one year later, Eanes and Glenner published their work about filamentous protein components of amyloid-laden tissue studied by X-ray diffraction procedures. They concluded from the X-ray diffraction patterns that the material forms a "cross-β" structure. They observed an intense ring at 4.75 Å, an overlaying diffuse halo at 4.3 Å and a broad and less intense ring at 9.8 Å.[54] These features were interpreted as a "pleated sheet" structure consisting of amyloid polypeptide chains regularly folded on themselves, where adjacent chain segments were laterally and antiparallel arranged.[54] The data also suggested that the axes of the chain segments run orthogonal to the filament axis.[54]

In the last 50 years, the progress in our understanding of the structure of amyloid aggregates has been remarkable and has mainly been driven by further X-ray and electron diffraction experiments as well as the rapid evolution of solid-state nuclear magnetic resonance (ssNMR) and cryo-electron microscopy (cryo-EM). For a detailed review on about half a century of amyloid research see Ke *et al.*, 2020.[55]

1.2 Molecular Structure of Protein or Peptide Amyloid Fibrils

With the report of cross-β fibrils,[53,54] a milestone in the general understanding of amyloid fibrils was reached. Nowadays, a commonly accepted biophysical definition of amyloid fibrils is that they show a cross-β fibre diffraction pattern when measured by X-ray diffraction.[56] The origin of these patterns is caused by the unique way the peptides or proteins are assembled. Elongated peptide or protein strands (β-strands) stack on top of each other forming β-sheets.[56] The β-strands within the β-sheets are held together by hydrogen bonds formed between their amide groups, [56–59] resulting in the measured distance of 4.8 Å of the β-strands within the β-sheets, Figure 1.2. The interaction of pairs of β-sheets to form a protofilament is supported by the interdigitating or butting of amino-acid side chains extending on both sides of the β-sheets. Because of the interaction of the side chains between the β-sheets, protofilaments are sometimes also called steric zippers.[56–59] The length of the side chains determines the distance of the two interacting β-sheets and is typically between 5 and 16 Å, Figure 1.2. The interfaces between the β-sheets are, with few exceptions, devoid of water.[58]

The self-assembly of these aggregates is energetically favourable and stable for several reasons. Water plays a significant role, specifically it orients near non-polar surfaces, which adversely affects its 3D hydrogen-bonding network.[60] This leads to a decrease in configurational entropy. Reducing the hydrophobic surface interacting with water leads to increased configurational entropy of water molecules.[60] Therefore, the hydrophobic effect is experienced as an attractive force between two non-polar objects if they are brought sufficiently close together.[60]

The arrangement of the β-strands in the β-sheets can either be in- or out-of-register. The former means that the β-strands are orthogonal to the β-sheet direction and out-of-register that the β-strands are out of square, Figure 1.3A.[59]

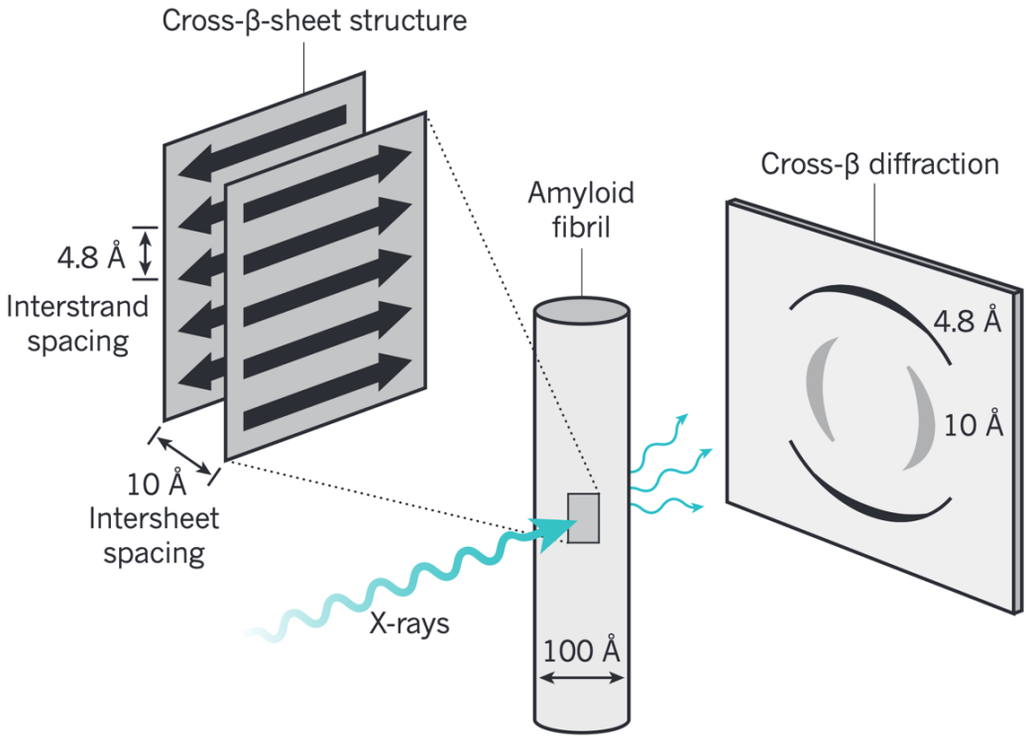


Figure 1.2. The cross-β-sheet motif.

A defining characteristic of amyloid fibrils is the typical X-ray fibre cross-β diffraction caused by their cross-β-sheet motif. The typical spacing of the β-strands causes a meridional reflection at about 4.8 Å and the distance between the stacked β-sheets causes a protein dependent equatorial reflection between 6-12 Å. In the Figure the distance of 10 Å is chosen as an example. Figure adapted from R. Riek & D. S. Eisenberg, 2016.[58]

If the β-strands are in-register in a β-sheet, energetically favourable side chain ladders can be formed. Typical ladders are formed of asparagine or tyrosine, Figure 1.3B.[59]

From theoretical considerations, it becomes clear that there are many possible β-strand arrangements. Assuming that the interacting β-strands are homo-steric, four distinctly different amyloid structures are adoptable by a single sequence. It is the case because each β-strand can be orientated, N- to C-terminus, and has a distinct front and back side. These two variables lead to two arrangements in which the β-strands are parallel and two in which the orientation of the β-strands alternates (antiparallel). In addition, these arrangements can be found with separated face and back sides (antifacial) or alternating between face and back side (equifacial), Figure 1.4.[59]

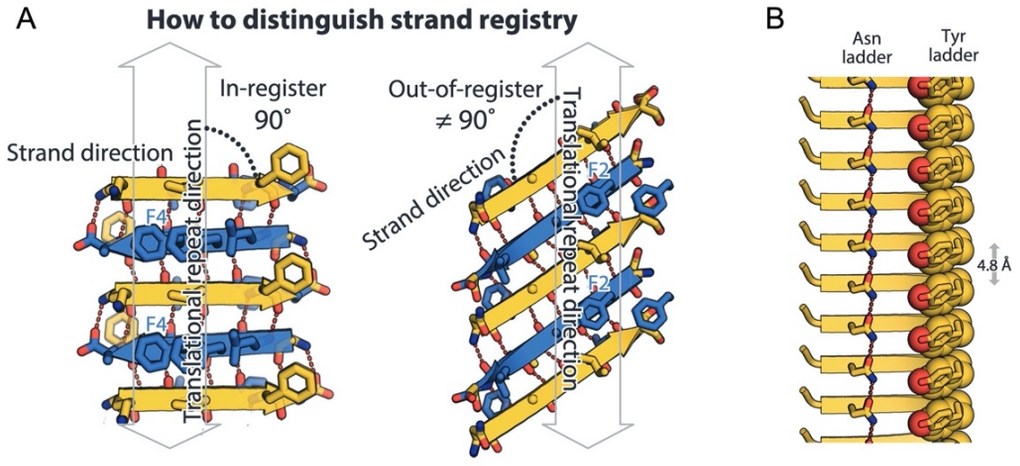


Figure 1.3. In- and out-of-register β-sheets and energetically favourable ladders.

A) In-register sheets are defined by strands orthogonal to its translational repeat direction (white vertical arrows) in comparison to out-of-register sheets. If the angle is not orthogonal, the sheet is out-of-register. B) In-register β-sheets are capable of forming energetically favourable ladders, for example, asparagine (Asn) or tyrosine (Tyr) ladders. Figure adapted from D. S. Eisenberg & M. R. Sawaya, 2017.[59]

These four different β-sheets can be arranged in ten different classes of homo-steric zipper amyloid spines, Figure 1.4. These different possible β-strand arrangements are biological significant as it means that a single sequence can theoretically self-assemble in many ways, which is known and observed as amyloid polymorphism.[59] This may be the molecular basis of prion strains.[61,62] The number of different steric zipper amyloid spines increases even further if segmental polymorphs, combinational polymorphs, intra-/intermolecular polymorphs and single-chain registration polymorphs are considered. For excellent reviews on the different structures of amyloid fibrils that have been observed, peptide fibrilization in general and the influence of peptide aggregation on current and future nanotechnology see D. S. Eisenberg & M.R Sawaya, 2016, I. Hamley, 2007 and Wei *et al.*, 2017.[59,63,64]

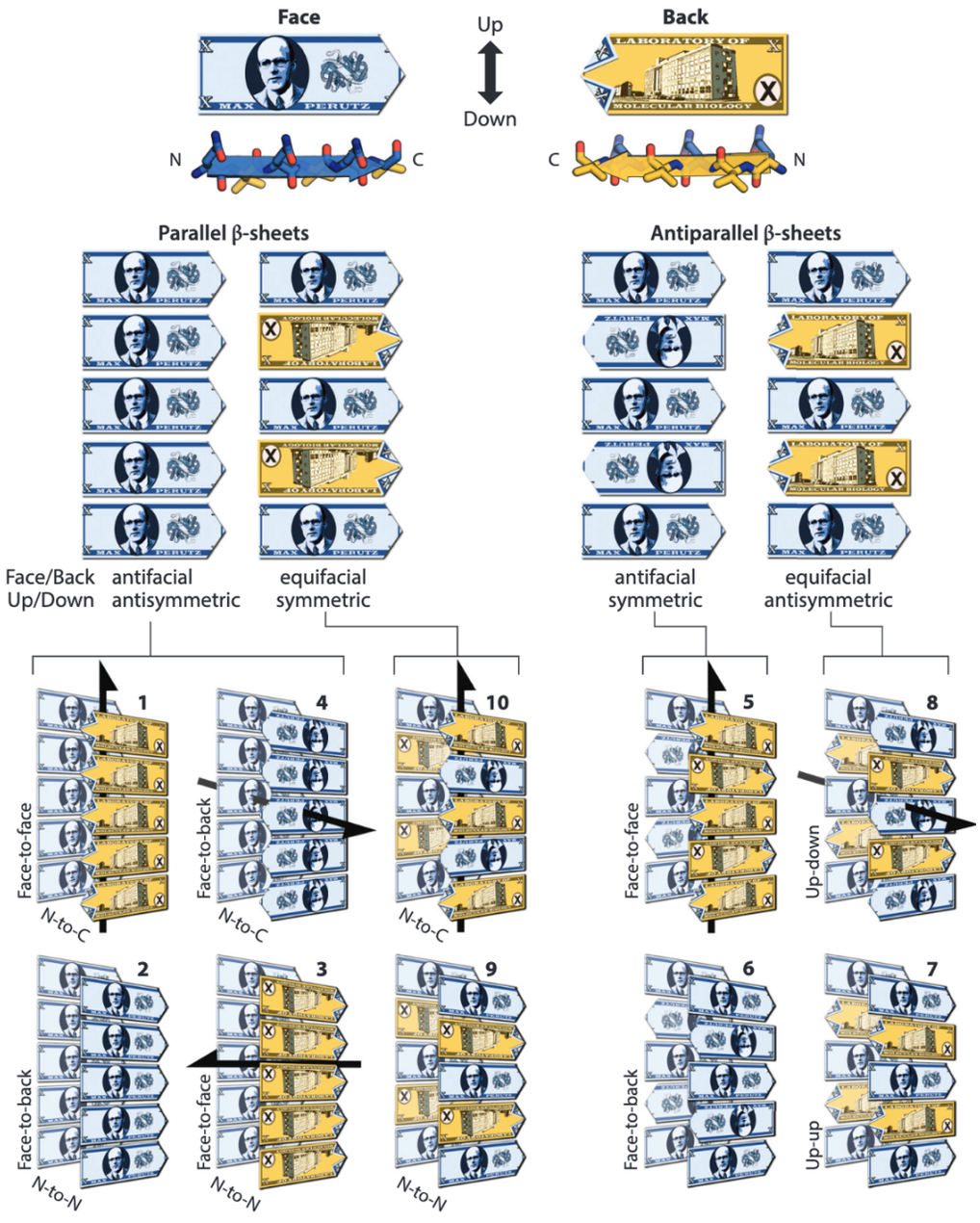


Figure 1.4. Symmetry classes of homo-steric zipper amyloid spines.

β-strands showing the C- to N-terminal direction, up/down edge and front/back face are illustrated as fictitious Max Perutz bank notes. They can be arranged in four different β-sheet conformations. These four different β-sheets can form ten different homo-steric zipper amyloid spines based on symmetry considerations, also called classes. Single-bladed arrows signify 2_1 symmetry axes, a 180° rotation about the arrow. Shifts of one-half or a whole repeat distance between β-strands of a β-sheet are possible. Figure adapted from D. S. Eisenberg & M. R. Sawaya, 2017.[59]

1.3 Atomic Structure of Glucagon Fibrils

An example of the atomic structure of a fibrillar aggregate of a peptide hormone is shown in Figure 1.5. The structure of glucagon at pH 2 was resolved by Gelenter *et al.* using ssNMR and ^{13}C and ^{15}N-labeled glucagon.

The observed length of the β-strand is formed by the complete length of the 29 residues of glucagon.[65] Such a long β-strand is remarkable as most commonly observed β-strands in known amyloid fibril structures and β-barrel membrane proteins only show β-strands containing fewer than thirteen amino acids.[66,67]

Glucagon aggregated at pH 2 also showed not only one but two different β-strand conformations. Conformer I and Conformer II can be differentiated from each other by their different wet and dry interfaces. Conformer I adopts a steric zipper structure with its odd-numbered residues, whereas Conformer II uses even-numbered residues. These two different conformers form the basic unit of glucagon fibrils at low pH which is a dimer of dimers.[65] The symmetry class can be described as antiparallel, antisymmetric and equifacial like class 8 in Figure 1.4. The structure of two molecular conformations integrated within a single ultrastructural fibril morphology is unique among the amyloid fibril structures known today, as well as containing the longest known β-strand among all peptides and proteins.[65]

The fibrillar structure of glucagon is a good example of typical side chain interactions. The steric zipper of glucagon formed at pH 2 is stabilised by a series of different side chain interactions. These favourable interactions include one tetrameric ‘aromatic box’ (F6-W25), several polar interactions (Q3-N28, T7-G24, S11-Q20, D15-S16), cation-π interactions (Y13-R18) and steric complementarity (M27-G4). The cation-π interaction mitigates an unfavourable electrostatic repulsion caused by the Arg18 side chains at the homodimer interface.[65]

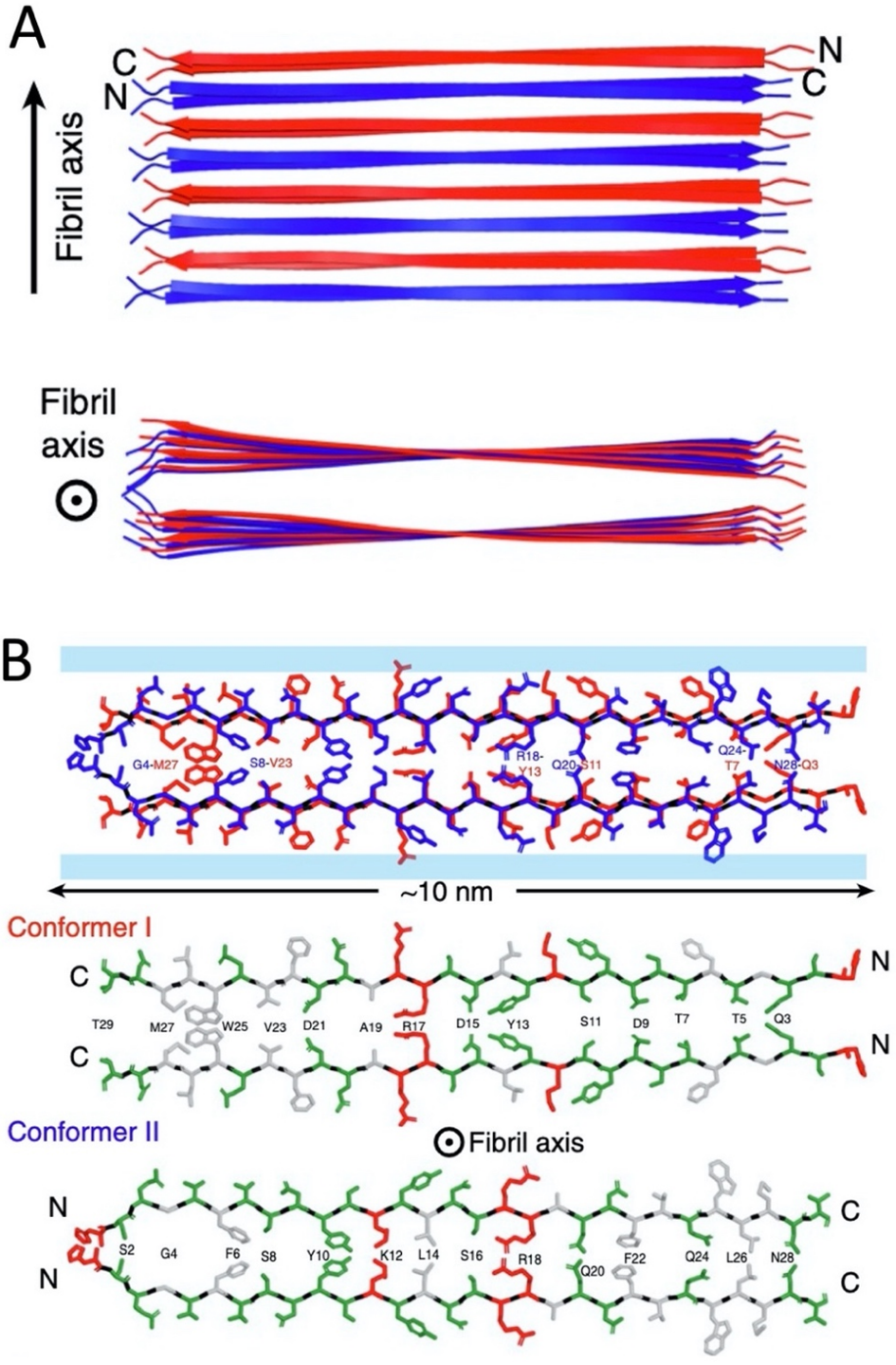


Figure 1.5. Atomic structure of a glucagon fibril at low pH.

Atomic-resolution structure of glucagon fibrils at pH 2 determined using ssNMR. A) Side and top view of the fibril. The conformation of the β-strands in the β-sheet alternate between Conformer I (red) and Conformer II (blue). It is similar to equifacial antisymmetric (class 8). However, one β-sheet is shifted by one position so that the β-strands are face-to-face. B) Side-chain packing in the dimer-of-dimer subunit steric zipper. The fibril axis is pointing out of the page. The light blue shaded region represents the solvent. Detailed image of the two conformers with hydrophobic side chains coloured white, polar residues green, and positively charged residues red. Figure adapted from Gelenter *et al*., 2019.[65]

Alterations in pH, temperature, mechanical conditions, or the peptide concentration can easily influence the interaction of amino-acid side chains. This leads to the assumption that the glucagon sequence can form a diverse set of 3D structures.[65] This is supported by the fact that at a pharmaceutical concentration of ~287 μM (~1 mg mL^{-1}) glucagon not only exists as straight fibrils but also in a twisted morphology. The system shows a peptide concentration dependence as the twisted species become more predominant below 72 μM (0.25 mg mL^{-1}). Additional fibrils morphologies can be formed by glucagon under different pH, temperature, and mechanical conditions.[9,68]

1.4 Aggregation to Form other Types of Structures

Higher-order structures comprised of amyloid-like fibrils have been observed for a number of proteins. Insulin and β-lactoglobulin have been shown to form spherulites, spherical structures consisting of radiating fibrils from a central core.[69] In addition, a number of peptides that self-assemble into fibrillar structures have also been shown to form hydrogels.[63,70,71] The formation of hydrogels can also be influenced by lipidation. The gut-derived peptide hormone PYY3-36 was shown to form hydrogels if lipidated but not in its unlipidated form.[34] The measured peptide concentration of the formed hydrogels was in the range of 3.5-5.5 mM.[34] For an excellent book chapter about colloidal gels from proteins and peptides see Djabourov *et al*., 2013.[72]

Amorphous aggregates, also known as disordered aggregates, are unstructured and contain no ordered intermolecular interactions.[73] They are often granular in appearance and can vary in size and solubility. When insoluble they form precipitates, which presents a major challenge for biotechnology.

1.5 Kinetics of Fibrils Formation

Aggregation is a complex phenomenon and can occur from many different states. In many cases, aggregation occurs from an unfolded or largely unstructured state, i.e., in the case of many peptides and intrinsically disordered proteins.[74–77] For amyloid fibril formation, the monomeric peptides first self-

associate to form oligomeric structures which go on to form a critical nucleus, Figure 1.6. Elongation of the nucleus is rapid relative to its formation giving rise to the sigmoidal aggregation kinetics often observed, Figure 1.6. The lag time caused by the primary nucleation process which can be shortened or eliminated by the addition of seeds.[78,79]

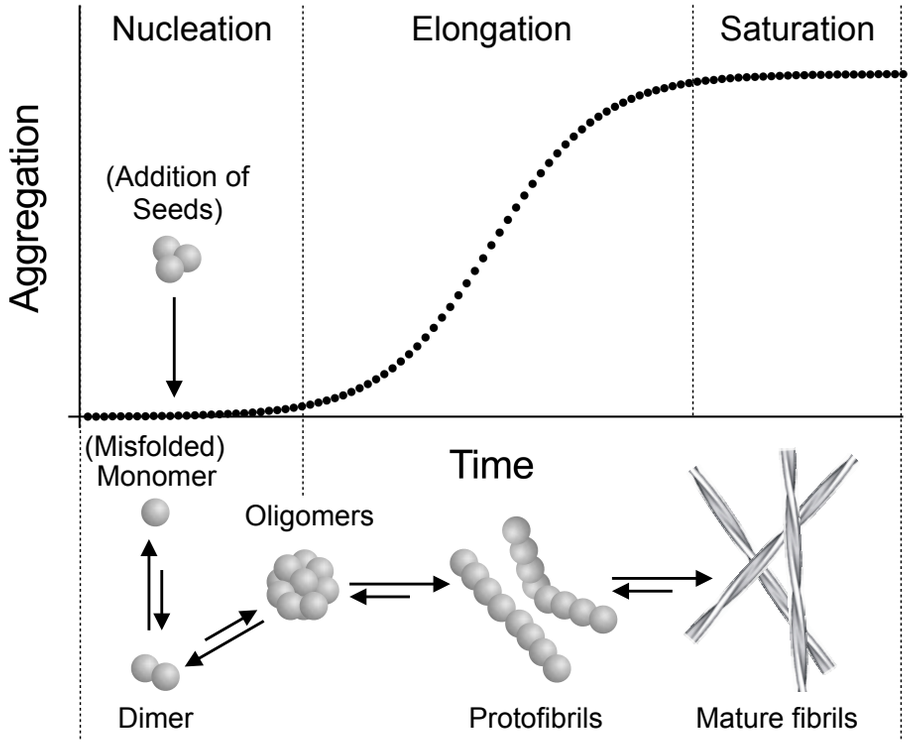


Figure 1.6. General mechanism of amyloid formation

Simplification of the mechanism of aggregation of amyloid fibrils following sigmoidal kinetics. Shown are the primary nucleation, elongation and saturation phases.

1.5.1 Different Aggregation Mechanisms

As shown in Figure 1.6, the kinetics of aggregation/fibril formation can often be described by a sigmoidal curve, with a lag phase, growth phase and a final plateau, at which point the system either completely depletes the pool of monomer, or the system reaches equilibrium between peptide monomers, oligomers and fibrils. Since the middle of the last century, significant research has been conducted in order to understand the aggregation process of polypeptide chains in detail.[80] In 1961, Oosawa and Kasai introduced a nucleated-polymerisation mechanism and a theory on the helical aggregation of macromolecules in which each amino-acid residue binds to three

neighbouring residues instead of one, as it would in a simple linear polymer.[81] Two decades later, Ferrone & co-workers added a double nucleation mechanism in which the fibrils autocatalyze aggregation by the nucleation of additional polymers at the surface of existing ones.[82] Recently, Cohen & co-workers described an analytical solution for the formation and growth of fibrils including not only primary nucleation and growth elongation but fragmentation of filaments and surface-catalysed secondary nucleation,[83] Figure 1.7A. Their approach has been successful in describing and predicting the aggregation behaviour of the Aβ42 peptide.[84,85] In their model, nucleation is a high-order reaction and oligomeric species that are formed prior to the final fibrillar structure are all on-pathway. In this case, lag times depend strongly on the peptide/protein concentration and lag times decrease with increasing peptide concentration.

1.5.2 Off-Pathway Oligomers or Off-Pathway Species

In the last ten years, there have been a number of observations on peptide and protein aggregation which can't be explained by these mechanisms alone. For example, it was shown that an amyloidogenic fragment of native human islet amyloid polypeptide (hIAPP)[86] and GLP-1(7-37)[5], under certain conditions, have aggregation kinetics that show no dependence on peptide concentration. Such unusual results are not consistent with a pure on-pathway oligomer mechanism and have been explained, in part, by the presence of off-pathway oligomeric species populated during the aggregation reaction. Rhoades & co-workers also introduced an off-pathway oligomer model for peptide aggregation in 2000 in which the monomeric peptide concentration is buffered by the existence of off-pathway micelles[87], Figure 1.7C. They proposed that above a minimum peptide or protein concentration, the minimum micelle concentration (CMC), off-pathway micelles are formed, Figure 1.7C. In this case, the formation of the micelles maintains the monomer concentration at the CMC. The formation of micelles is one property peptides have in common with surfactants and detergents. Supporting this, it was also shown that Aβ peptides lower the surface tension of water with increasing peptide concentration similar to the behaviour of surfactants.[88]

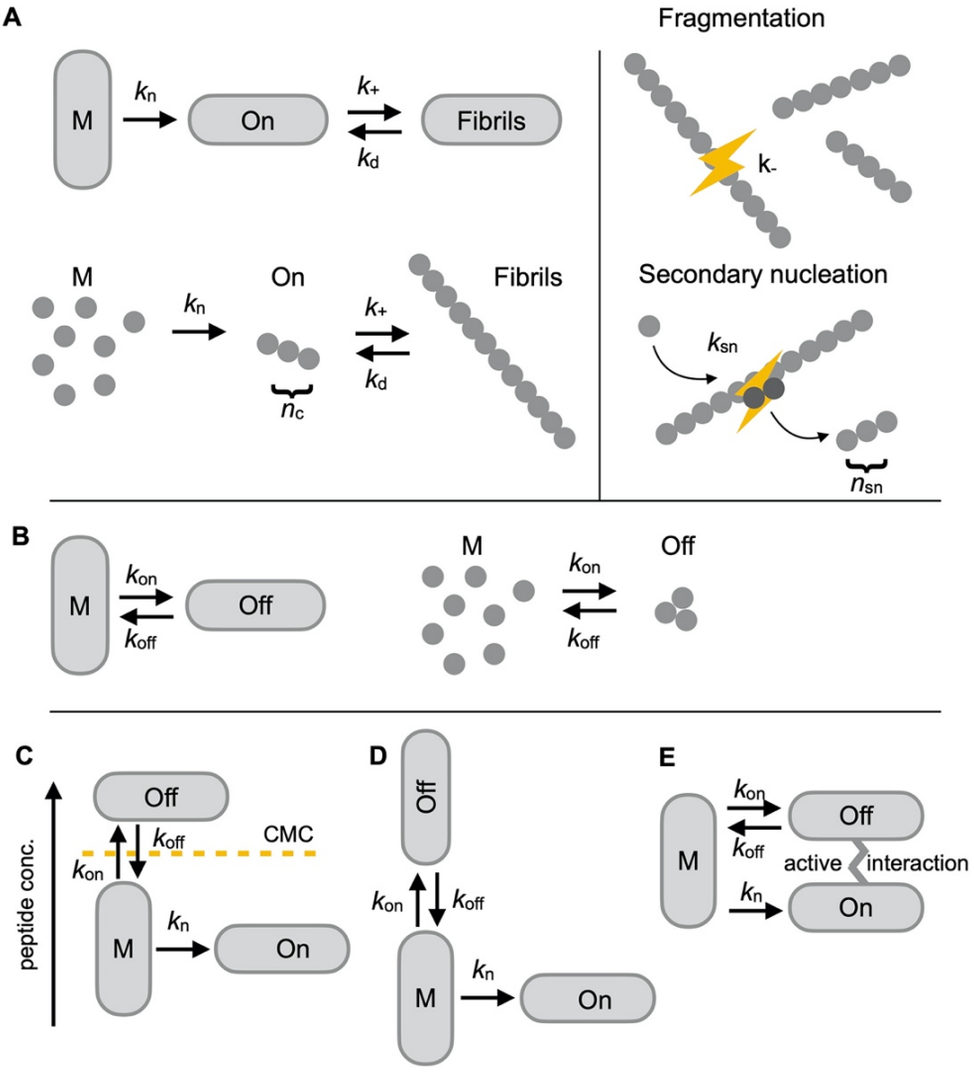


Figure 1.7. Different mechanisms involving on- and off-pathway oligomeric species.

A) Simple on-pathway model and additional secondary nucleation where M is monomer, on are on-pathway oligomers, and off are off-pathway oligomers. Primary nucleation with the primary nucleation rate k_n leads to on-pathway nuclei consisting out of n_c monomers. They are assumed to be stable and no dissociation happens. Fibrils are formed by elongation or through fragmentation and by monomer-dependent secondary nucleation. The rate constants for elongation, dissociation, fragmentation and secondary nucleation are k_+, k_d, k_- and k_{sn}, respectively. In a nucleation-elongation model, a higher monomer concentration M leads to faster aggregation. B) Simple off-pathway model with the rate constants k_{on} and k_{off} as association and dissociation rate constants, respectively. C) On- and off-pathway species mechanism following a critical peptide concentration (CMC) model. At the CMC the system is saturated and excess monomers associate as off-pathway aggregates. D) On- and off-pathway species model without a CMC. The two pathways are independent of each other. E) On- and off-pathway species model in which there is an interaction between the different species. This can lead to aggregation reactions in which there is an increase in lag time and $t_{1/2}$ with increasing peptide concentration. This could be caused by a change in the fluidity of the solution with increasing peptide concentration, or the binding of off-pathway species to binding sites of on-pathway aggregates.

In 2008 it was shown by Powers and Powers[89] that aggregation mechanisms in which a parallel off-pathway process occurs additional to the on-pathway process can not only lead to constant t_{lag} and $t_{1/2}$ but also to increasing $t_{1/2}$ values with increasing peptide/protein concentration. In their model, the increase of $t_{1/2}$ with increasing peptide/protein concentration can be explained by the effect of the off-pathway species which buffer the available monomer concentration during the aggregation process, Figure 1.7D. Subsequently the rate of fibril formation is restricted but the amount of peptide/protein available for the formation of fibrils is not.

However, t_{lag} itself cannot increase with increasing peptide/protein concentration, as shown mathematically in Section 4.8. The observation of increasing t_{lag} with increasing peptide concentration can only be explained by more complex models, for example, one in which there is an active interaction of the off-pathway species with on-pathway species (Figure 1.7D), in comparison to simple on-pathway models or other on- and off-pathway models.[90]

1.6 Factors Affecting the Physical Stability (Aggregation) of Peptides

It is well established that the physical stability of peptides is affected by many factors including temperature, excipients, freeze-thaw, processing, peptide sequence, peptide concentration, pH and chemical modifications.[73,91–95] A brief summary of the different factors affecting the physical stability of peptides follows.

1.6.1 Amino-acid Sequence

How the amino-acid sequence of a peptide affects its intrinsic propensity to aggregate, to form either amorphous or amyloid-like fibrils, is now very well understood, such that there are numerous prediction programmes available.[73,96] Factors such as hydrophobicity, charge state, β-sheet forming propensity along with other properties are all known to contribute.[73,96–98] A consecutive series of five or more residues which all have a high intrinsic propensity form what is

known as an aggregation-prone region, APR.[96,99–102] In some cases, these methods have been used to predict the effect of an amino-acid substitution on aggregation propensity, for example, the effect of pathogenic mutations on the aggregation of Aβ *in vitro* and *in vivo*.[96] A typical example of a prediction of structural aggregation using the TANGO method is shown in Figure 1.8.

A number of approaches have been developed to predict aggregation propensity some of which are primarily based on our knowledge of the intrinsic propensity of amino acids to aggregate and form fibrils. As our understanding of the factors that affect aggregation in terms of amino-acid sequence has improved, methods have also been developed that not only take into account APRs in peptide sequence but also other factors such as the presence of gatekeeper residues that help prevent aggregation,[103–105] and the flexibility and solvent accessibility of the peptide chain.[106] Other algorithms also take into account the propensity of amino acids to form other forms of stable structure such as α-helices, as aggregation is always in competition with the formation of other conformational states.[107]

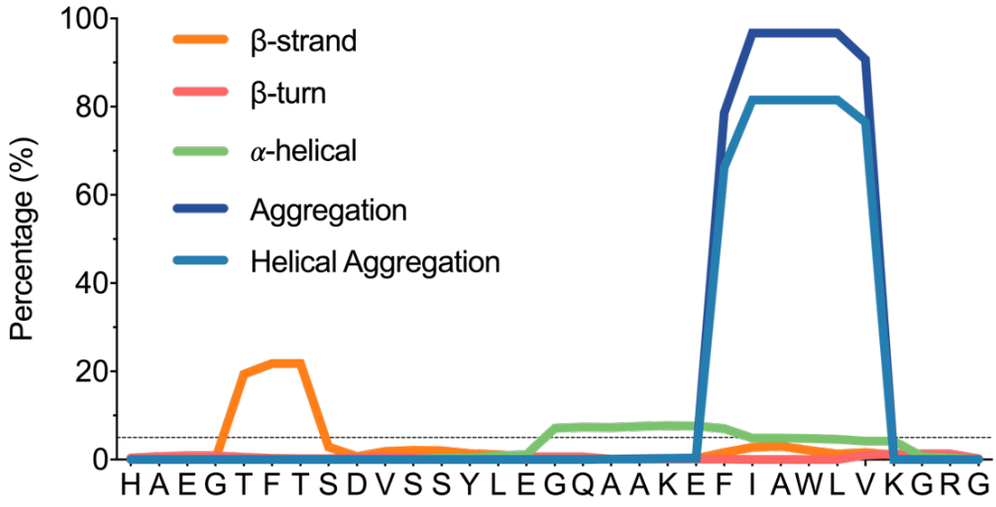


Figure 1.8. Prediction of structural aggregation.

Aggregation propensity of GLP-1 predicted by the Tango algorithm at pH 7.5. Any segment with an aggregation tendency above 5% over five or more residues in length is a potential aggregating segment. The data are shown as percentage of β-strand (orange line), β-turns (red line), α-helical (green line), aggregation (dark blue line) and helical aggregation (bright blue line).

Since high-resolution structures of amyloid fibrils have become available from X-ray crystallographic and NMR studies,[58] structure-based methods for predicting APRs have also been developed.[73,96] In these cases, APRs which are compatible with the formation of the cross β-structure seen in the core of amyloid fibrils are identified by threading peptide and protein sequences through templates of different amyloid structures scoring each resultant structure in terms of packing density,[108] pairing energies for side chains that pack against each other in the structure[109–111] and other metrics.[112,113] In another case, a combination of the two approaches has been used and a position-specific scoring matrix developed taking into account both intrinsic propensities and also structure.[114]

Name	**Description of method**
Zygreggator	APRs and gatekeepers, protein flexibility and solvent accessibility[106]
TANGO	statistical thermodynamics approach. Identifies APRs but takes into account competition β-aggregated state and other conformations[107]
SALSA	uses β-propensity for each amino acid and a sliding average to identify APRs[102]
AGGRESCAN	uses aggregation propensity for each amino acid[115,116]
FoldAmyloid	packing density[108]
3D profile method	structure-based method uses amyloid fibril structure as a template[112]
Pre-AMYL	structure-based method uses amyloid fibril structure as a template[113]
PASTA	pairing energies calculated for pairs of residues within an amyloid structure. Sequences are scored by energies to identify APRs[109,110,117]
BETASCAN	β-strand pairing energies using amyloid core structure[111]

Table 1.2. Methods for predicting aggregation-prone regions and aggregation propensity in peptides.

For peptides which do not form stable tertiary or quaternary structures, these approaches work well, and predictions are generally reasonably accurate. In addition, while many algorithms which are not structure-based predict both the propensity for a peptide to form either an amorphous or highly structured aggregate, the structure-based methods are largely accurate only for predicting the formation of amyloid-like fibrillar structures.

Given these prediction methods, summarized in Table 1.2, it is relatively straightforward to identify APRs in largely unstructured peptides and to rationally design amino-acid substitutions to reduce aggregation propensity. Of course, such changes may also affect activity. At the moment, there are no programs available to predict the effects of incorporating non-natural amino acids nor the effects of chemical modification of a peptide on its aggregation propensity.

1.6.2 Peptide Concentration

One of the most important factors influencing the physical stability of peptides is peptide concentration. This is long-established through numerous kinetic studies on the formation of amyloid fibrils,[118–122] amorphous aggregates[123,124] and adsorption on surfaces.[125] With respect to the formation of amyloid-like fibrils, extensive kinetic studies have been performed over wide ranges of peptide concentrations and, in the majority of cases, the kinetics have been shown to be consistent with a nucleation – polymerization mechanism, Figure 1.7A.[126] In this case, the monomeric starting material is converted into oligomers, and ultimately a species that can act as a nucleus for rapid fibril growth. Addition of usually monomers to a nucleus or growing fibrils occurs rapidly to form protofibrils and mature fibrils.

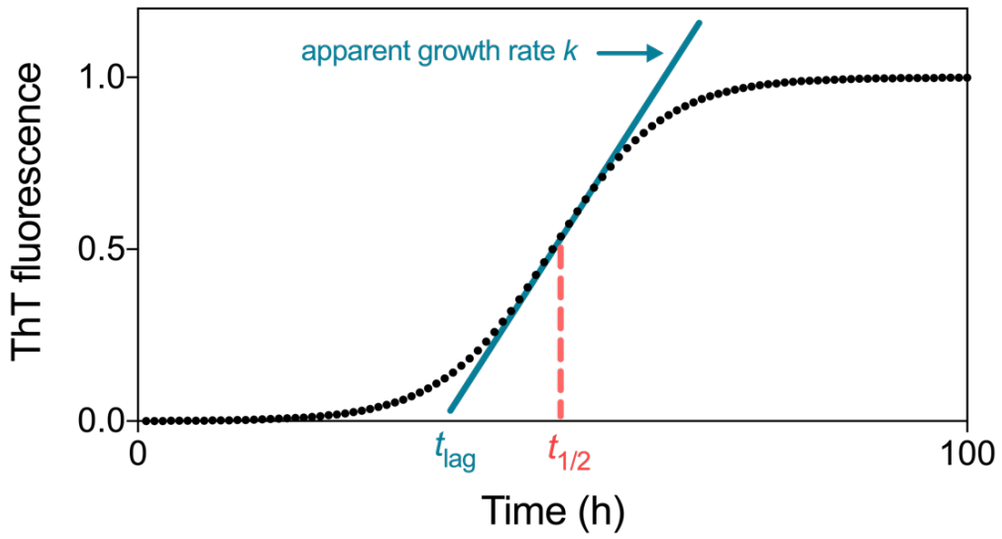


Figure 1.9. Definition of key kinetic parameters obtained from sigmoidal aggregation kinetics.

The curve shows the key kinetic parameters that can be obtained from a fit of the aggregation data to a simple sigmoidal function: $t_{1/2}$, the time at which the ThT fluorescence is 50% of its final value (red dashed line), the apparent growth rate which corresponds to the gradient of the steepest part of the rapid growth phase (blue line), and t_{lag}, the lag time at which point sufficient nuclei are present in solution for elongation and growth of fibrils to become rapid.

The essential features of such systems are that the kinetics show a lag phase, a rapid elongation or growth phase and a plateau. During the lag phase, a number of oligomeric species, including nucleating species, and short amyloid fibrils form until a point is reached at which the growth of fibrils is rapid, Figure 1.9. The formation of fibrils then slows as a plateau is reached which represents either the point where monomers are completely depleted or an equilibrium point at which mature fibrils are in equilibrium with starting monomers and other species. The sigmoidal kinetic traces can be fitted to obtain key kinetic parameters such as lag time, $t_{1/2}$ and the apparent growth rate, Figure 1.9. How these kinetic parameters depend upon peptide concentration has been established for many systems including Aβ[127], α-synuclein[128], tau[129], glucagon-like peptide-1 (GLP-1)[5] and other peptides. The majority of peptides and conditions show kinetics typical of nucleation – polymerization mechanisms, and in these cases, generally both $t_{1/2}$ and t_{lag} decrease with increasing peptide concentration. The growth rate has been shown to decrease or increase, or be independent of peptide concentration, depending upon the peptide/conditions and rate-limiting step. Cases where there is an increase in growth rate with

peptide concentration point towards association of monomer to the fibril being rate limiting to the growth. In contrast, if the growth rate is independent of peptide concentration this can indicate that some other step, such as a conformational change of the monomeric unit, is rate limiting. The growth rate is also dependent upon secondary nucleation steps, if they exist, such as fragmentation of the fibrils to form small fibrils with a net increase in reactive fibril ends from which growth can occur, or secondary nucleation sites on the surface of a fibril, Figure 1.7A.[119]

The fact that $t_{1/2}$ and t_{lag} decrease with increasing peptide concentration highlights the issues of physical stability of peptides, which may not be a significant issue at low peptide concentrations, but which can become a serious problem at high concentrations such as those frequently used in therapeutics.

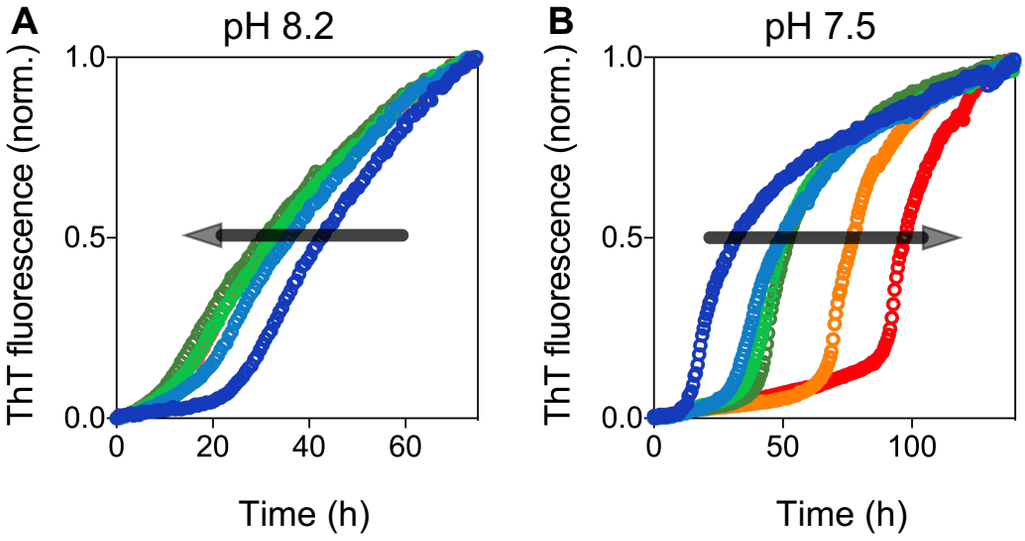


Figure 1.10. Dependence of GLP-1 aggregation kinetics on peptide concentration at two different pH values.

Data adapted and normalised from Zapadka *et al.*[5] GLP-1 incubated at six different concentrations (dark blue circles 25 µM, bright blue circles 50 µM, bright green circles 75 µM, dark green circles 100 µM, orange circles 125 µM, red circles 150 µM) and at 37 °C. A) Aggregation kinetics of GLP-1 in Tris buffer at pH 8.2 with decreasing $t_{1/2}$ and t_{lag} with increasing peptide concentration. Most likely following a simple nucleation – polymerization mechanism with only on-pathway oligomeric intermediate states. B) Aggregation kinetics of GLP-1 in phosphate buffer at pH 7.5 where $t_{1/2}$ and t_{lag} increase with increasing peptide concentration.

In stark contrast to the systems and kinetics described above, there are an increasing number of peptides and conditions under which very different kinetics have been observed. In these cases, increasing concentrations of peptide leads to no change or an increase in $t_{1/2}$ and t_{lag}.[5,89,130] This unusual kinetic behaviour has been attributed to the formation of a different form of oligomeric species where the oligomers formed are off-pathway, and either cannot directly be converted into on-pathway oligomers or nucleating species, or the interconversion is extremely slow, or where the off-pathway oligomers even actively suppress the on-pathway aggregation, Figure 1.7C,D & E. One such system is GLP-1, which is a particularly interesting system as, in this case, there is a switch in the kinetic behaviour with pH.[5] At pH > 8.0, GLP-1 follows the more common nucleation – polymerization kinetics with $t_{1/2}$ and t_{lag} decreasing with increasing peptide concentration, Figures 1.10A and 1.11A & C; however, at pH 7.5, the kinetics reverse such that both lag time and $t_{1/2}$ increase with increasing peptide concentration indicating off-pathway oligomers have formed,[5] Figures 1.10B and 1.11B & D.

At an intermediate pH, 8.0, there is little dependence on peptide concentration as there is a delicate balance between on- and off-pathway species.[5] Although off-pathway oligomers have been observed now for a number of peptides and proteins,[5,89,130] these systems have not been studied in the detail that peptides and proteins showing simpler nucleation – polymerization kinetics have been and therefore not much is known about the nature of off-pathway oligomers, for example, their size, structure and cellular toxicity.

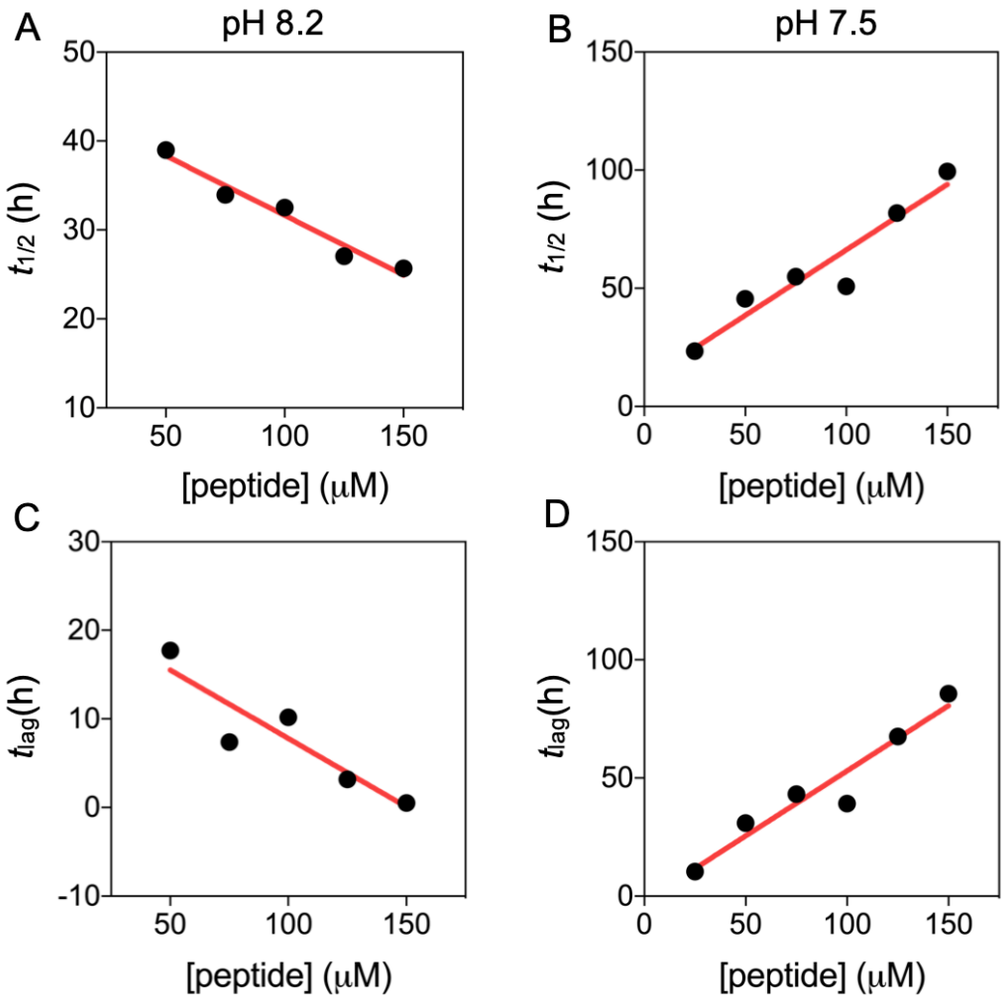


Figure 1.11. Influence of peptide concentration on GLP-1 aggregation parameters $t_{1/2}$ and t_{lag} at different pH values.

Data adapted and normalised from Zapadka et al.[5] GLP-1 incubated at 25, 50, 75, 100, 125 and 150 μM and at 37 °C. A) Dependence of $t_{1/2}$ on peptide concentration for the aggregation of GLP-1 into amyloid fibrils at pH 8.2 under conditions where the peptide follows a simple nucleation–polymerization mechanism with on-pathway oligomers only. B) Dependence of $t_{1/2}$ on peptide concentration for the aggregation of GLP-1 into amyloid fibrils at pH 7.5. C) Dependence of t_{lag} on peptide concentration for the aggregation of GLP-1 into amyloid fibrils at pH 8.2 under conditions where the peptide follows a simple nucleation-polymerization mechanism with on-pathway oligomers only. D) Dependence of t_{lag} on peptide concentration for the aggregation of GLP-1 into amyloid fibrils at pH 7.5 under conditions where on- and off-pathway oligomers are populated.

1.6.3 pH and Net Charge

Electrostatic interactions are known to play major roles in the self-association of peptides to form all types of aggregates.[131–136] Essentially, for systems carrying an overall net charge, unfavourable electrostatic repulsion needs to be overcome in order for self-association/aggregation to take place. Generally, it has been shown for many systems, the higher the net charge the slower the aggregation and the lower the net charge the higher the propensity to aggregate.[131–136] As such, solution conditions such as pH, ionic strength, nature of cations and anions in solution and the presence of polyelectrolytes can all have a dramatic effect on both the rate and extent of aggregation. One of the best characterized systems in terms of the effects of net charge, ionic strength, and counter ions on aggregation is the 37-residue positively charged peptide IAPP which has no acidic groups.[137] A comprehensive and rigorous study has shown that there are multiple effects on aggregation depending upon the pH.[137] At pH 8.0 there is a strong dependence on ionic strength due to Debye screening and increasing ionic strength accelerates aggregation by up to a factor of ten. In this case, specific ion effects were also observed, the nature of the cation being relatively unimportant but the rate depending crucially on the identity of the anion. For example, at pH 8.0 and a total ionic strength of 30 m, the $t_{1/2}$ measured in sulphate and fluoride differed by a factor of four. The effects were observed both on the lag phase and the apparent growth phase. At pH 5.5, where IAPP has a higher overall net charge, the aggregation kinetics were considerably slower than at pH 8.0, but there were similar effects of ionic strength.[137] For lysozyme and acylphosphatase, studies have established that polyelectrolytes can also have significant effects. In these cases, a number of different polyelectrolytes were found to bind to and stabilize mature fibrils.[138]

Although it is easy to establish some general rules of thumb with regards to the effects of pH, net charge, ionic strength, ions and polyelectrolytes, these rules should be viewed with caution. There are an increasing number of examples where these rules are not followed due to the complexity of aggregation reactions. For example, around the isoelectric point of peptides and proteins, where net charges are essentially zero, then monodisperse and quasi-amorphous aggregates have been observed instead of amyloid-like fibrils.[69]

Thus, aggregation kinetics can be very different from those of amyloid-like fibril formation which is often seen at pH values away from the isoelectric point. In addition, other changes in mechanism have been found at different values of pH and net charge.

For example, as already discussed in Section 1.6.2, the aggregation kinetics of GLP-1 into amyloid-like fibrils is highly sensitive to pH.[5] In this case, the presence of off-pathway species slows the rate of aggregation considerably, particularly at high peptide concentrations. Thus, although the net charge on the protein at pH 7.5 is lower than at pH > 8.0, the rate of aggregation is considerably slower, contrary to what would be expected.[5]

In previous studies, GLP-1 aggregation showed only minor changes with ionic strength. It was shown that 50, 75, 100, 125 and 150 μM GLP-1 in 25 mM citrate buffer at pH 3.5 show no change in aggregation kinetics at 0, 50 or 150 mM NaCl (ionic strength of the samples 18, 68 and 168 mM, respectively).[139]

1.6.4 Chemical Degradation

Chemical degradation pathways in biological pharmaceuticals have been reviewed many times.[140] A number of chemical degradation pathways have been characterized including: deamidation and isomerization,[141] oxidation,[140,142] hydrolysis, disulfide bond breakage and formation, succinimidation,[143] deglycosylation, Maillard reaction, racemization and β-elimination and many of these have been shown to increase the aggregation of biologics.[143] This is in effect because chemical degradation/modification often changes the physical properties of a peptide/protein such as hydrophobicity, secondary structure (relevant for relatively unstructured peptides) and/or tertiary structure, and the thermodynamic and/or kinetic barriers to unfold (relevant only for highly structured proteins and not discussed further here).[143]

1.6.5 Surfaces and Interfaces

Surface-induced aggregation, often called ‘surface adsorption’, is a physical degradation process whereby there is accumulation and adhesion of peptide molecules to a surface. During this process, peptide molecules change their

physical state, i.e., they are no longer in solution.[144–147] There are different mechanisms of surface adsorption of biomolecules and some of the major factors driving adsorption include intra-molecular forces, hydrophobicity,[148–153] and ionic or electrostatic interactions.[147,154]

Extensive research has been conducted on the surface adsorption of human insulin.[150,155–159] It is well known that insulin aggregates to form fibres in the presence of hydrophobic surfaces[149,150,157,158] and especially in a peristaltic pump system.[159] In this case, the potential formation of aggregation-prone intermediates is very problematic as this may catalyse the aggregation of other molecules in bulk solution. Other studies have established that the aggregation of $A\beta_{1-40}$ into fibrils is also influenced by surfaces.[160–164] Surface adsorption has been studied for other systems such as GLP-1,[154] GLP-2,[165] acylated glucagon-like peptide-2,[165,166] glucagon,[167] poly(ethylene glycol) (PEG)-glucagon[168] and α-synuclein.[151,169]

It has been demonstrated that any interface (between solid, liquid or air) will affect the ability of water molecules to form a dynamic hydrogen-bonding network as in bulk water, and therefore air–water and ice–water interfaces are described in the literature as hydrophobic interfaces, because these interfaces do not allow hydrogen bonding.[146] Adsorption of peptides through their hydrophobic side chains to such hydrophobic interfaces can thus 'hide' the hydrophobic surface from the aqueous bulk. Some research shows that, for example, $A\beta_{1-40}$[161,170] and insulin both aggregate at air–water interfaces.[171]

Many studies have confirmed that the rate of peptide and protein adsorption is normally governed by diffusion and is thus a function of concentration.[143] The amount of surface-adsorbed species slowly increases over time, usually due to further peptide aggregation at the site.[160] Of course, there may also be slow structural rearrangements of the adsorbed peptide.[160] The adsorbed layer often changes its properties over time; this is associated with a change from a reversible to an irreversible process.[147] The final morphologies of the aggregates formed can be sensitive to the surface chemistry.[154,156,160] For example, charged surfaces promote protofibril formation, while hydrophobic surfaces promote formation of spherical amorphous clusters.[160]

It is well known that after sufficient time most biomolecules will adsorb to solid

surfaces so strongly that they will not (or only very slowly) desorb from the attached surface back into solution.[172] However, it is possible to some degree to control this process by adding surface active agents which can lead to desorption. The degree of adsorption depends on multiple factors, including concentration, pH, excipients and temperature.[146,147]

1.6.6 Excipients

Excipients of many types have been employed to try and reduce aggregation in numerous systems including salts, surfactants, osmolytes, preservatives, chelators and antioxidants, specific ligands, sugars and carbohydrates. Below is a brief description of some of the excipients that have been used and their effects on aggregation.

Buffering agents

As already discussed, pH strongly influences peptide stability and the potential for biomolecular aggregation. The most commonly used buffers in pharmaceutical development are acetate, citrate, histidine, phosphate, Tris and glycine.[93,143] A complete description of the effects of buffer on peptide aggregation is beyond the scope of this chapter; however, there is an excellent review by Zbacnik and co-workers.[173]

Salts

Standard physiological salts have been used as tonicity modifiers in peptide formulations.[93,143] Tonicity measures the effective osmotic pressure gradient of two solutions separated by a semipermeable cell membrane. It differs from osmotic pressure by taking only solutes into account, which cannot cross the membrane.[174] Salts have complex effects on the physical stability of biomolecules affecting both conformational and colloidal stability. Their effects frequently vary according to the surface charge on the peptide or protein, and the overall effect of a salt on physical stability is a balance of different and multiple mechanisms by which salt interacts with water and biomolecules. Salts can influence physical stability by altering the properties of the peptide–solvent system (Hofmeister effects) and by screening electro-

static interactions (Debye–Hückel effects). Many different studies on the dependence of amyloid formation on ionic strength over the past ten years have shown that ions can influence both the kinetics of aggregation and the structure of fibrils formed. This was, for example, shown for islet amyloid polypeptide (37-residue),[137] human glucagon (29-residue),[175] and Aβ peptide.[134,136,176]

Surfactants

Surface-active agents are amphiphilic molecules that tend to orient such that the exposure of the hydrophobic part of the molecule to the aqueous solution is minimized. Non-ionic surfactants such as Tween20 and Tween80 are often added to peptide solutions during pharmaceutical development to prevent aggregation or adsorption during purification, filtration, shaking and transportation.[3,93,177–179] However, the chemical stability of Tween20 and Tween80 in pharmaceutical formulations is very important. Unfortunately, both surfactants are known to undergo oxidation and cleavage at their ethylene oxide subunits, as well as hydrolysis of the fatty acid ester bond.[180–183] Thus, their use can be complicated as these chemical processes can potentially promote precipitation in some peptide formulations. The use of Tween80 is also far from straightforward as it has been reported that it can have a dual effect on physical stability. For example, Tween80 in the formulation of the protein IL-2 mutein inhibited shaking-induced aggregation, but also had a dramatic effect on the oxidation and aggregation of the protein during storage of the liquid formulation.[184,185] To counteract this issue, methionine and tryptophan are used in formulation development to prevent Tween from oxidative degradation.[186] There are two excellent reviews on the interaction between surfactants and peptides, see Otzen[187] and Khan and co-workers.[178]

Amino acids

A number of free amino acids have been used to stabilize proteins and reduce aggregation. Arginine, histidine, lysine, glycine and aspartic acid have been found to reduce aggregation for a number of biomolecules.[93]

Osmolytes

Osmolytes such as sucrose, trehalose, sorbitol and glycine are commonly used in formulation development.[93] Sucrose has been shown to inhibit IL-1ra dimer formation.[143] However, their use is not always beneficial and it has been reported that some osmolytes accelerate the rate of aggregation, for example, of glucagon.[188]

Antioxidants and chelators

Peptide oxidation is a major cause of chemical instability and also sometimes linked to physical stability. Amino acids such as methionine, cysteine, histidine, tyrosine and tryptophan in peptides are susceptible to oxidation under some conditions encountered during pharmaceutical development. Therefore, a number of antioxidants are used as excipients including ascorbic acid.[93] It has also been reported that sodium thiosulfate, methionine, catalase or platinum, and the chelating agents EDTA and DTPA are also effective in reducing the oxidation of biologics.[93,189–191]

Preservatives

Preservatives are used as antimicrobial agents and are very common in liquid formulations.[93,192] Their role is to prevent bacterial growth during storage.[93] The most commonly used preservatives in pharmaceutical development are *m*-cresol, phenol and benzyl alcohol. However, it has been reported that preservatives can sometimes cause protein aggregation. For example, benzyl alcohol induces aggregation of recombinant human interleukin-1 receptor,[193] and *m*-cresol induces aggregation of cytochrome *c*.[194] Moreover, destabilizing effects of phenol and *m*-cresol were reported for insulin lispro and insulin aspart.[195] Little is known about their effects on the physical stability of peptides.

Polymers and Proteins

A number of polymers have been shown to stabilize drug products including polyethylene glycols (PEG). The process of both non-covalent and covalent attachment or amalgamation of PEG polymer chains is described as PEGylation. Other polymers that are used as excipients include dextrans,

heparin, gelatins type A, poly-L-glutamic acid, poly-L-lysine, poly-Asp and poly-Gly.[93,143] For example, it has been reported that different PEGs stabilize the proteins lysozyme and bovine serum albumin.[143] Human serum albumin is also commonly used as an excipient to inhibit protein/peptide adsorption onto surfaces during product development.[143] However, adding protein-based excipients adds more complexity to pharmaceutical formulations.[93]

Recently, some excellent reviews on biologics formulation[196,197] and excipients, including protein–excipients interactions[93,197,198] have been published.

1.6.7 Impurities

The results of *in vitro* aggregation studies are often controversial and difficult to reproduce by others. One of the reasons for this problem is the variable amounts of intrinsic impurities in peptide and protein preparations that can affect aggregation rates. For example, the aggregation kinetics of synthetic Aβ has been reported to vary not only batch-to-batch but also with storage and solubilization conditions.[199] In line with these findings, kinetics studies revealed that the impurities left in the synthetic peptides of glucagon decreased fibrillation rates dramatically compared to that of pharmaceutical-grade glucagon.[68,175,200]

1.6.8 External Factors

Numerous other external factors including temperature, pressure, agitation and lyophilization/freeze drying can affect the aggregation of peptides, in many cases, in ways difficult to predict.

Pressure

High hydrostatic pressure (HHP) has been used in a number of studies.[201] It is now well established that HHP pushes the conformational equilibria towards ensembles of structures that occupy smaller volumes, it acting on cavities that are excluded from solvent that can be found in the hydrophobic cores of natively folded proteins and aggregates including fibrils.[201] HHP has been used

to dissociate multimers in oligomeric structures and, in some cases, dissociate aggregates.[201,202] The degree to which HHP acts to dissociate aggregated forms of peptides depends upon the structures that are formed and also their inherent thermodynamic stability. For example, if fibrils have cavities within them (cavity defects as they are called) then often low pressures are sufficient to dissociate the aggregated form; for example, early aggregates and protofibrils are generally more readily dissociated by HHP than mature fibrils.[203] For a peptide from transthyretin that can form fibrils, $TTR_{105–115}$, 220 MPa dissociates aggregates and protofibrils while mature fibrils remain stable up to 1.3 GPa.[203] For α-synuclein, pathogenic variant fibrils are less resistant to HHP than those of the wild-type protein, hence more toxic oligomers are present.[201] The dissociation of aggregates and fibrils by HHP is reversible and can be enhanced by the addition of cosolvents such as glycerol, methanol and dimethylsulfoxide.[203]

Temperature

Aggregation kinetics frequently show non-Arrhenius behaviour even over short temperature ranges.[204] A consequence of this is that it is challenging to extrapolate the results of aggregation kinetics commonly measured at higher temperatures to lower temperatures. In addition, cold denaturation of amyloid fibrils has been observed,[205,206] and low temperatures have been combined with high pressures to dissociate aggregates and fibrils.[203] Experiments at increased temperatures between 50 and 100 °C with globular proteins has shown heat-induced gelation can occur for some systems.[207]

Agitation

Agitation, in the form of a magnetic stirring bar or by use of a shaker table, is routinely used in studies of peptide and protein aggregation. It has been employed in order to accelerate otherwise slow aggregation kinetics and also used to improve the reproducibility of experimental results.[208] Despite the frequent use of agitation little is known about the shear forces that are produced and the effects on different steps in the aggregation reaction. In addition, as stirring/shaking is often different in different studies, it is challenging to

undertake a comparative analysis on systems studied in different research groups.[209]

Lyophilization/freeze drying

The stabilization of peptides in a solid form is a very common approach to increase both their chemical and physical stability.[210] Lyophilization and freeze drying are frequently used to prepare peptides in a solid state. However, for a number of systems studies have shown these processes can result in conformational changes leading to increased aggregation after reconstitution.[211] These conformational changes can be caused by different effects such as stress induced by formation of ice-water interfaces, freezing induced pH-changes of specific buffers, changes in concentration of solutes and phase separations.[146,212]

Parameters that can be controlled to minimise the effect of freeze-drying are the choice of the buffer,[213] the duration of the freezing process[214] or additives, which minimise the interaction of the peptide with the water-ice interface, for example, surfactants.[146]

1.6.9 Chemical Modifications

Chemical modifications of different types are increasingly used to optimize the chemical/physical and biological properties of therapeutic peptides. This includes the incorporation of non-natural amino acids,[215] engineering of disulfide bonds[216–220] and lipidation or acylation of the peptide.[221,222]

Lipidation/acylation of peptides is becoming increasingly important in the field of therapeutic peptides.[223,224] Inspired by endogenous lipidation processes, it has been shown to improve many properties of a peptide, particularly its ability to bind to human serum albumin increasing the half-life of the peptide *in vivo* dramatically and also enhancing its ability to permeate body tissues.[223,224] Current therapeutics which use this approach include Liraglutide[225] (a lipidated form of GLP-1), and insulin detemir, a lipidated conjugate of insulin.[226]

Although considerable research has been undertaken on how this class of chemical modification affects the activity of the peptide *in vitro* and *in vivo*,

rather little is known about how these modifications affect the physical stability of the system. One recent biophysical study on Liraglutide by Wang and co-workers showed that this lipidated peptide can form oligomers in solution.[227] In this case, a pH-sensitive equilibrium between octamers and dodecamers was observed in the vicinity of pH 7.[227] The importance and impact of oligomers, synthetic processes and lyophilisation on the aggregation of Liraglutide was also shown by Bothe *et al.*[91] Different oligomeric states showed different aggregation kinetics as well as different potencies. They were also able to show that Liraglutide can maintain its oligomerisation state in dried lyophilised powders, acting as a memory effect from its synthetic process and purification.

In another case, acylation of the anti-microbial peptide novicidin resulted in an increased tendency to form α-helical structure when a C_{12}-acyl group was attached. The peptide concentration dependence of this conformational change was attributed to the increased tendency of the peptide to form micelles as shown by NMR.[228] Although not peptides with any therapeutic potential, the Prive group has published extensively on the properties of lipopeptide detergents (LPDs) which have been developed for structural studies of membrane proteins.[229] LPDs, which consist of an α-helical peptide with alkyl chains at either end of the helix, have been shown to self-assemble into cylindrical micelles that have a densely packed hydrophobic core.[229]

In contrast to the chemical modifications used by the pharmaceutical industry to optimize the properties of therapeutic peptides, where little is known about the effect of the modifications on physical stability, a great deal has been reported for peptides whose aggregation is associated with disease states. In these cases, the modifications investigated are largely post-translational modifications that occur *in vivo* and which, in many cases, are thought to affect the disease state. For example, the effects of many post-translational modifications on the aggregation of Aβ peptides have been studied. An N-terminal truncation and pyroglutamate modification at residue 3 of Aβ peptide has been shown to have a higher propensity for oligomerization and aggregation than full-length Aβ.[230,231] In addition, phosphorylation of Aβ has been shown to increase oligomer formation,[232,233] while citrullination of Aβ decreases the rate of fibril formation but may increase the amount of toxic

oligomers.[234] Similarly, nitration of tyrosine 10 (Y10) of Aβ has also been shown to increase the amount of oligomers of Aβ in solution.[235] Backbone modification of Aβ at Gly29 also shows decreased fibril formation and increased oligomer formation.[236] Similarly, differences in the sequence in the form of mutations influence the aggregation pathway and, therefore, the progression of Alzheimer diseases. Common mutations of Aβ are A21G Flemish, E22K Italian, E22Q Dutch, E22G artic, and D23N Iowa. More details from a Chemist's perspective about Aβ, its mutations and its role in Alzheimer's Disease and fibrillation can be found in an excellent review by I. W. Hamley.[38]

The aggregation and fibril formation of α-synuclein is associated with another neurodegenerative disease, Parkinson's Disease. α-synuclein is a 14 kDa intrinsically disordered protein.[237] It is known to consist of three main domains: An amphipathic N-terminal involved in membrane binding,[238] an aggregation-prone hydrophobic central region,[239] and an acidic unstructured C-terminal region.[238] Different autosomal dominantly inherited point mutations (Ala53Thr, Ala30Pro, and Glu46Lys) have been described.[240–242] They all can cause Parkinsonism and, as shown for Glu46Lys, also result in clinical and pathological features characteristic of Lewy body dementia.[242] There is also a large body of data that establishes that numerous post-translational modifications *in vivo* affect the propensity of the protein to aggregate and thus potentially the disease state. For example, N-acetylation has been shown to decrease fibrillation,[243,244] post-translational modifications due to oxidative stress such as HNE conjugation, nitration and oxidation have been shown to reduce fibril formation but increase oligomer formation,[245,246] and oxidation of Met5 also decreases fibril formation but increases the concentration of oligomers.[247]

The intracellular protein tau, associated with Alzheimer's disease, is also known to undergo a large number of post-translational modifications *in vivo* including phosphorylation, ubiquitination, nitration, truncation, prolyl isomerization, glycosylation and glycation. Many of these are thought to affect neurofibrillar tangle formation, the aggregated form of tau in cells.[248] A few, but not many, of these post-translational modifications on tau aggregation have

been studied in detail. For example, cysteine guanylation of tau has been shown to inhibit fibril formation with a corresponding increase in the population of oligomers.[248]

The effects of a large number of other modifications on the physical stability of specific peptides or proteins have also been reported, including different glycosylations[249] and fatty acid conjugation.[250]

1.7 Lipidation of Peptides

The focus of this Thesis is the physical stability of GLP-1, GLP-1 with amidated C-terminus and two lipidated forms of GLP-1 with amidated C-termini. The idea of using lipidation to influence the properties of peptides *in vivo* is directly motivated by examples observed in Nature, so called endogenous lipidation.

1.7.1 Endogenous Lipidation

Lipidation of polypeptides can be found naturally in mammalian cells, for example, palmitoylation or myristoylation. S-palmitoylation, in particular, is a reversible covalent post-translational attachment of palmitic acid to the amino acid cysteine. It is the most common acylation of polypeptides in eukaryotic cells and can be found in many proteins including G-proteins, ion channels, receptors, cytoskeletal proteins and kinases.[222,251] The modification affects, amongst other things, the regulation of protein subcellular localisation, stability, trafficking, translocation to lipid rafts, aggregation and the interaction with effectors.[222] Palmitoylation is reversible and dynamically regulated by enzymes.[222]

Several naturally occurring acylations of different proteins are presented in Table 1.3. The spectrum of lipids is broad, ranging from caprylic acid, through geranylgeranyl groups to cholesterol. The attachment can be catalysed by enzymes, can be non-enzymatic, autocatalytic or the mechanism is still unknown. In the case of GAP-43, a heterogeneity in the naturally attached fatty acid was found with both C_{16} or C_{18} fatty acids observed, showing that either palmitoyl transferase can attach different fatty acids as substrates or there exist

other transferases which are, as yet, unknown.[222,251]

Attached lipid	Modified amino acid	Enzyme that catalyses the attachment	Modified protein
Octanoic acid	Serine	Ghrelin O-acyl transferase (GOAT)	Ghrelin
Myristic acid	N-terminal Glycine	N-Myristoyl transferase	Recoverin
Palmitic acid	Cysteine	Palmitoyl transferase	PSD-95
Palmitoleic acid	Serine	Porcupine	Wnt3a
Steeric acid	Cysteine	Unknown	GAP-43
Farnesyl group	Cysteine	Farnesyl transferase	H-Ras
Geranylgeranyl group	Cysteine	Geranylgeranyl transferase I	RhoA
Palmitic acid	N-terminal Cysteine	Hedgehog acyl transferase (Hhat)	Hedgehog
10-nitro-9,12-*cis*-octadecadienoic acid	Cysteine	Non-enzymatic	Keap1
Cholesterol	Carboxylate group of a carboxy-terminal Glycine	Autocatalytic	Hedgehog

Table 1.3. Examples of endogenous protein lipidation.[222]

1.7.2 Influence of Lipidation in Therapeutics

Artificial lipidation can be used in a similar manner to that of Nature to improve the therapeutic effect of polypeptides.[221] Connected by covalent bonding, a lipid group (cholesterol, palmitic acid or farnesyl) can have a significant influence on the efficiency of a drug.[221] The lipid group can be attached to an alcohol, amino or acid group on a peptide/peptidomimetic inhibitor *via* amide, ether or carbamate linkages. Through its interaction with the membrane in cells, the lipid group has a strong influence on the location of the drug.

In Figure 1.12A, B and C, three different applications of lipidated peptides /proteins are shown. When anchored in a membrane, the lipidated peptide/protein can bind to soluble proteins or membrane-bound proteins. The former can be helpful in reducing the amount of a certain protein and the latter can take advantage of the higher molarity of membrane-binding peptides within the membrane.[252]

Lipidation can also influence cell-permeability as it was shown for palmitoylated peptidyl inhibitors by Tarasova *et al.*,[253] Figure 1.12D. However, the effect of lipidation is not restricted to its influence on the interaction with membranes. It also influences the binding affinity to plasma proteins as human albumin.[254] A schematic representation of a lipidated peptide that binds to serum albumin is shown in Figure 1.12E.

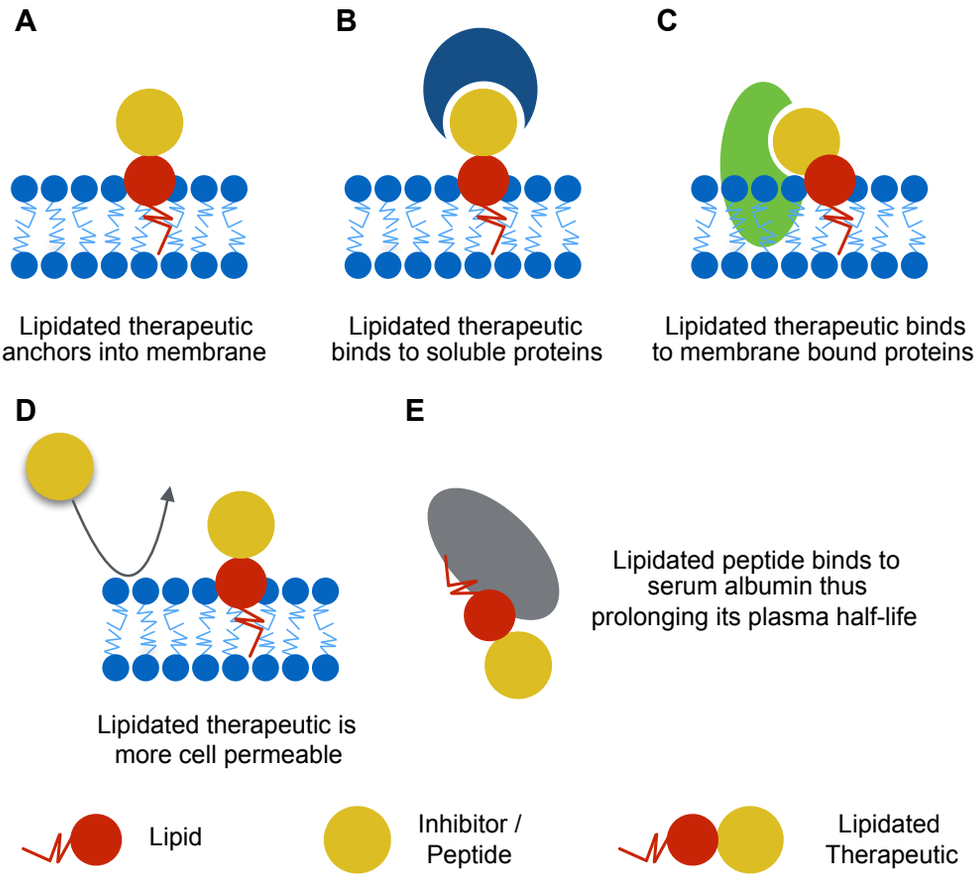


Figure 1.12. Different applications of lipidated peptides/proteins in vivo.

A) A lipidated inhibitor binds to a membrane. B) Soluble protein (blue moon-shaped molecule) immobilised by binding to the lipidated therapeutic. C) Membrane bound protein (green molecule) immobilised by binding to the lipidated therapeutic. D) Increased cell permeability due to lipidation. E) Lipidated peptide binds to serum albumin (grey molecule) preventing its degradation.

1.8 Glucagon-Like Peptide 1 (GLP-1)

The peptide used in these studies is glucagon-like peptide 1 (GLP-1 (7-37)). It is an incretin hormone and different modified variants are currently already in use as drugs to treat type two diabetes mellitus.[224,255,256] Previously, Zapadka *et al*. have investigated and reported how the peptide concentration dependence of the aggregation kinetics of GLP-1 (7-37) varies from pH 7.5 to 8.2.[5] The main finding was that there was a striking change in kinetic behaviour over this very short pH range, which was attributed to the population of off-pathway oligomeric species at pH values close to the pI of GLP-1. Values of $t_{1/2}$ and the lag time were found to decrease with increasing peptide concentration at pH 8.2 as expected for a nucleated-polymerisation mechanism. However, at pH 7.7, little effect of peptide concentration on these kinetic parameters was observed and at pH 7.5, they were found to increase with increasing peptide concentration.[5]

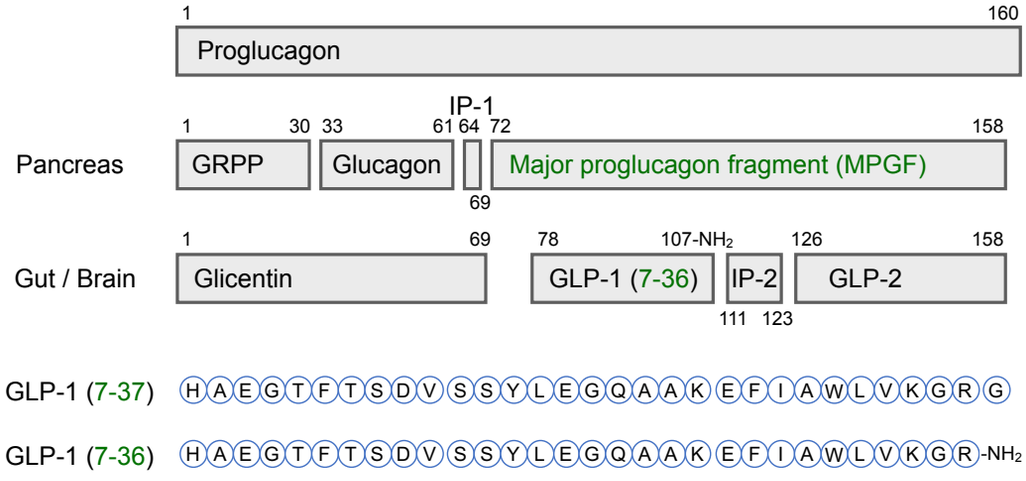


Figure 1.13. Products during the process of proglucagon.

The numbering of GLP-1 (7-37) and (7-36) refers to the major proglucagon fragment MPGF (green font colour).

GLP-1 is important in mammalian metabolism and plays a main role in preventing hyperglycemia. GLP-1 was described for the first time in 1986 as a specific fragment of proglucagon. Proglucagon itself was described by K. Lund and J Habener, who were able to deduce the structure of proglucagon of angler fish.[257] Bell and co-workers published the full human proglucagon structure in 1983.[258]

In addition to glucagon itself, proglucagon contains two glucagon-like structures. Framed by consensus cleavage sites for the prohormone convertase, the fragments for the two glucagon-like structures are proglucagon 72-108 and 126-158.[258] However, neither directly affected pancreatic insulin secretion, and only become active on truncation. This led to the conclusion that the sequence is processed in the pancreas as well as the gut. Holst and co-workers found that in the pancreas, the major proglucagon fragment (MPGF) is formed and secreted, whereas GLP-1 and GLP-2 were both formed and released in the gut.[259,260] The highly immunoreactive GLP-1 turned out to be a truncated form of the predicted GLP-1. Instead of proglucagon 72-107, the active sequence starts with residue 78, Figure 1.13.[261]

The naturally occurring and active human GLP-1 peptides are nowadays often referred to as GLP-1 (7-36) amide or GLP-1 (7-37) because of the original assumption that GLP-1 would consist of 37 amino acids. *In vivo*, the C-terminal Gly of GLP-1 7-37 can be amidated, Figure 1.13.[261,262]

1.8.1 Biological Activity of GLP-1

GLP-1 is, as mentioned, a highly immunoreactive incretin.[263] It can influence the blood sugar level in two different ways, first by stimulating glucose-induced insulin release and second by inhibiting glucagon secretion.[264] Both effects can limit hepatic glucose production, which is closely correlated to fasting hyperglycemia of type 2 diabetes. GLP-1 is also known to be a physiological and potent inhibitor of gastrointestinal secretion, both gastric and pancreatic, as well as motility.[265] Glucagon-like peptide 1 (GLP-1) is an intestinal hormone signalling nutritional abundance, with an unusual therapeutic potential. An overview over the physiology of GLP-1 secretion and its influence on different organs and tissue is given in Figure 1.14.

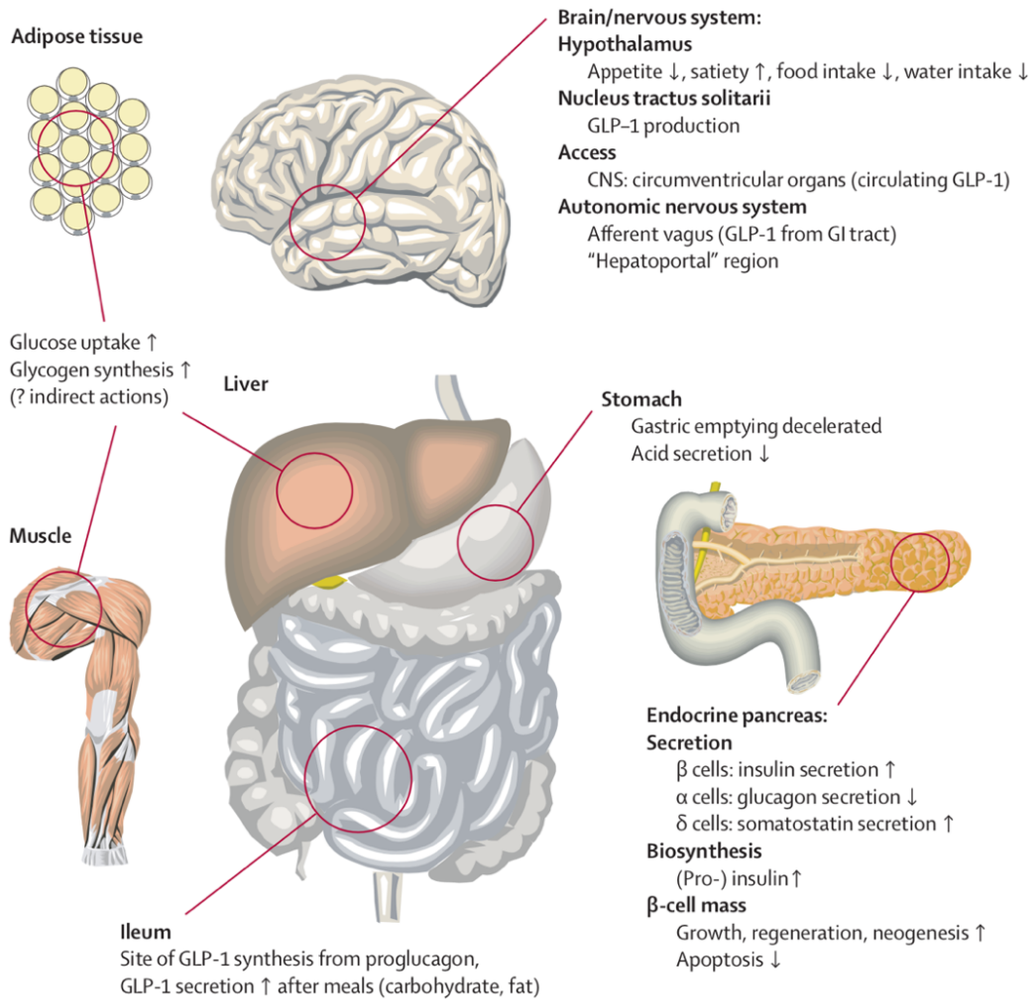


Figure 1.14. Physiology of GLP-1 secretion and action on GLP-1 receptors in different organs and tissues.

Figure adapted from D. Drucker & M. Nauck, 2006.[266]

1.8.2 GLP-1 as a Therapeutic Agent and its Challenges

The biggest challenge in the development of GLP-1 as a therapeutic was that it lasts for extremely short times in blood plasma. The measured half-life of GLP-1 in humans is around 2 min after injection. One major factor of its fast degradation is the enzyme dipeptidyl peptidase 4 (DPP-4).[267] Ahren and co-workers showed in 2004 that the use of DPP-4 inhibitors can prolong the half-life of GLP-1.[268] Another approach to minimise degradation uses the dependence of DPP-4 on the presence of a penultimate Pro or Ala at the N-terminus. It was shown that GLP-1 analogues, with Ser, Val or α-aminobutyric acid substituted at the relevant positions, stayed biologically active and were more resistant against degradation.[269] Unfortunately, the degradation through DPP-4 is only one factor in its degradation processes, and the half-life increased

only slightly up to 4-5 min.[269] The remaining degradation is caused by extraction of GLP-1 in the kidneys[270] although degradation by other enzymes may also play a role. It all meant that to overcome the challenge of a short half-life, more complex modifications were needed.

1.8.3 Modifications and Analogues of GLP-1

Strategies to overcome the short half-life of GLP-1 were developed using two different approaches. The first one uses a non-human peptide, isolated from the saliva of a reptile, the Gila Monster (*Heloderma suspectum*).[271] The 39-amino acid long peptide Exendin-4 has 53% homology to human GLP-1. Being a full and potent agonist of the mammalian GLP-1 receptor, Exendin-4 has the advantage that it is not sensitive to DPP-4 as well as the fact that it is barely extracted in the kidneys. Its half-life is about 30 min after injection and this can be increased further by subcutaneous injection, and adequate concentrations can be achieved for up to five hours.[272] This was sufficient for the development of a therapeutic based on two daily injections. Subsequently, a synthetic replica was produced and marketed under the name Exenatide. It was the first GLP-1 receptor agonist on the market in 2005.[272] A slightly modified version is Lixisenatide which is designed for once-daily use. Despite its improved usability, it is identical to Exenatide in terms of its pharmacokinetics.[273]

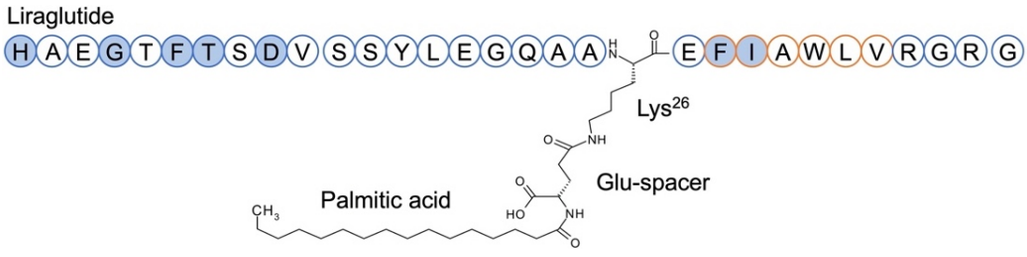


Figure 1.15. Sequence and structure of Liraglutide and its lipidation.

Blue highlighted residues are known to be important for binding to the receptor. Residues by Tango predicted to be important for amyloid formation are highlighted by orange circles.

The second approach was based on modification of human GLP-1 with lipids. The potency of different lipidated GLP-1 analogues was determined using a functional assay with cloned human GLP-1 receptors which were expressed in baby hamster kidney cells.[255] In general, lipidation at the carboxyl terminus (C-terminus) with long chain fatty acids did not affect the potency of the GLP-1. From earlier studies it was already known that the amino acids at position 7, 10, 12, 13, 15 in the N-terminal region and 28, 29 in the C-terminal region are important for binding and activation of the receptor as shown in Figure 1.15.[274] For the lipidation itself, Knudsen *et al*. used the amino acid lysine as the attachment site.

The first successful approach, known as Liraglutide, was released in 2009/10. It is a lipidated form of GLP-1. The attachment of palmitic acid *via* a glutamic acid linker at residue 26 (Lys) results in the ability to bind non-covalently to serum albumin. The latter increases its half-life in blood plasma by reducing the renal filtration as well as by improved protection from DPP-4 degradation. With a half-life of about 12 hours, Liraglutide is suited for a once-daily injection. It has also been approved to treat obesity.[275]

Figure 1.16. GLP-1 and GLP-1 receptor agonists used as pharmaceuticals.

The sequence of GLP-1 (7-37) (blue circles) in comparison to different GLP-1 receptor agonists. Red circles highlight modifications of the original sequence of GLP-1 (7-37).

To reach the aim of improved usability, several once-weekly use agonists have also been developed. This was achieved by coupling the peptide to larger molecules, for example human albumin or an IgG4 Fc domain, by covalent (Albiglutide[276–278], Dulaglutide[278]) or non-covalent (Semaglutide[279]) binding, Figure 1.16. The attachment to a larger molecule leads to slower clearance and improved half-life.

The future of GLP-1 receptor agonist (GLP-1RA) treatment stays an exciting topic, more so as patents of the first marked products will run out in the early 2020s.

1.9 Fatty Acids Used for Lipidation

A commonly used lipid for lipidation of GLP-1 are fatty acids, a carboxylic acid with a long aliphatic chain. Typical examples of saturated fatty acids are shown in Table 1.4. In this Thesis, all modifications of GLP-1 are based on palmitic acid. Palmitic acid is the most common saturated fatty acid found in animals, plants and microorganisms and is a significant component of palm oil, as the name suggests. The lipid is attached to the peptide, *via* a glutamic acid spacer to increase the distance to the peptide, Figure 1.15.

The lipidation of GLP-1 with palmitic acid and a glutamic spacer was based on detailed studies evaluating the influence of the position as well as different spacers and fatty acids with different lengths. Residues known to be important for binding to the receptor were excluded as lipidation sites.[274] Different fatty acids, spacers and positions were screened by Knudsen *et al.* (2000).[255] The combination of good biological activity and physical stability lead ultimately to the choice of glutamic acid and palmitic acid.

Carbon atoms	Sytematic name	Common name	Melting point (°C)
12:0	n-dodecanoic	Lauric	44
14:0	n-tetradecanoic	Myristic	54
16:0	n-hexadecanoic	Palmitic	63
18:0	n-octadecanoic	Stearic	70
20:0	n-eicosanoic	Arachidic	77

Table 1.4. Examples of saturated fatty acids.

1.10 Aggregation of GLP-1 under Acidic, Neutral and Basic Conditions

In previous studies, the aggregation of GLP-1 into amyloid fibrils was investigated in some detail at pH values between pH 7.5- 8.2.[5] Dr Karolina Zapadka extended these studies in previously unpublished work to characterise the physical stability of GLP-1 at low pH (3.0 and 3.5). Standard ThT assays were used to measure the aggregation kinetics to obtain the lag time, $t_{1/2}$ and apparent growth rate as a function of peptide concentration, Figures 1.17 - 1.21. At both pH 3.0 and 3.5, GLP-1 aggregates into β-sheet rich fibrils that bind ThT and which show sigmoidal kinetic traces typical of a nucleated polymerisation reaction. The timescales of aggregation at pH 3.0 and 3.5 are similar to those at pH 7.5-8.2, lag times varying from 5-50 hours depending upon peptide concentration, Figure 1.18A & B. Figure 1.18C & D show the results at pH 7.5 and 8.2, which are consistent with previously reported aggregation kinetics of GLP-1.

The dependence of lag time and $t_{1/2}$ on the initial concentration of monomeric peptide is shown in Figures 1.18, 1.19 and 1.20, and this provides important information on the mechanism of fibril formation. At pH 3.0, $t_{1/2}$ and lag time decrease with increasing peptide concentration, the slope of the plot of log peptide concentration vs log lag time, γ, is -1.0 ± 0.2, Figure 1.20A. This suggests that, at this pH, only on-pathway oligomers play a major role and that the filamentous growth reactions are limited by primary and surface-catalysed secondary nucleation processes rather than by simple breakage.[280] These results are similar to those obtained at pH 8.2, Figure 1.17D.[5] In contrast, at pH 3.5, both the lag time and $t_{1/2}$ increase with increasing peptide concentration similar to the results at pH 7.5, Figures 1.17B, 1.18B and 1.19B. These results suggest, under these conditions, off-pathway oligomeric species are present and significantly affect the aggregation kinetics.

Additionally, at pH 3.5 and 7.5, there is an increase in ThT fluorescence in the initial baseline of the aggregation assays, Figure 1.17B & C. This indicates either that small fibrils are being formed or that other species such as on- or

off-pathway oligomers capable of binding and increasing the ThT signal are populated. However, the formation of early fibrils is unlikely as secondary nucleation mechanisms should result in an immediate rapid growth phase instead of the observed approx. relatively slow linear increase seen during the lag phase. Therefore, it is probable that the off-pathway oligomers have some capacity to bind to ThT and increase its fluorescence.

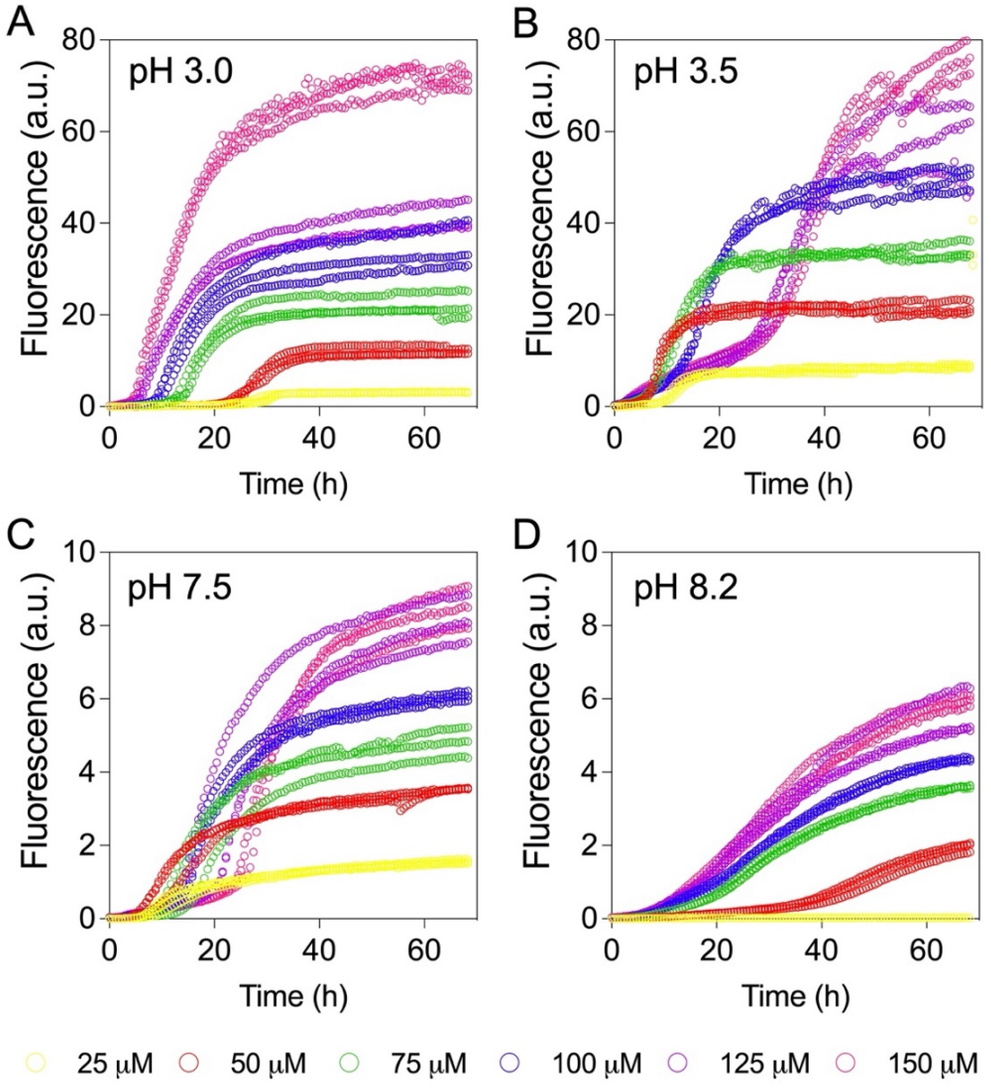


Figure 1.17. Aggregation kinetics of GLP-1 as a function of peptide concentration at different pHs.

Data collected by Dr Karolina Zapadka. GLP-1 aggregation kinetics in 25 mM citrate buffer at pH 3.0 (A), 25 mM citrate buffer at pH 3.5 (B), 25 mM phosphate buffer at pH 7.5 (C), and 25 mM Tris buffer at pH 8.2 (D). Panels A-D show typical traces for the fibrillation of GLP-1 followed by ThT fluorescence at different peptide concentrations: 25 µM (yellow), 50 µM (red), 75 µM (green), 100 µM (blue), 125 µM (purple), and 150 µM (pink). Each peptide concentration was performed in triplicate.

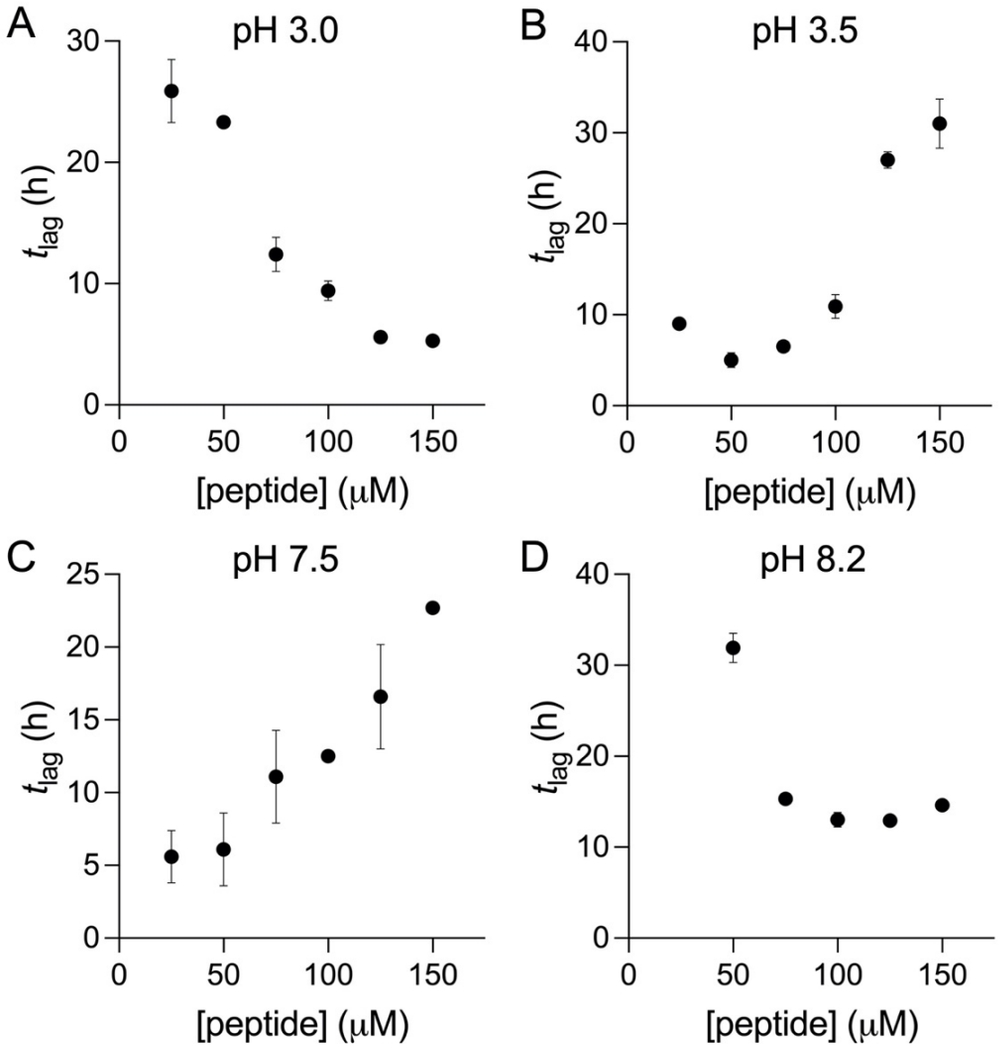


Figure 1.18. Dependence of the GLP-1 aggregation lag time on the peptide concentration.

Data collected by Dr Karolina Zapadka. Dependence of the lag time on the concentration of GLP-1 in 25 mM citrate buffer at pH 3.0 (A), 25 mM citrate buffer at pH 3.5 (B), 25 mM phosphate buffer at pH 7.5 (C), and 25 mM Tris buffer at pH 8.2 (D). All samples incubated at 37 °C. The error bars shown in A-D correspond to the standard deviations of the kinetic parameters determined from one experiment where each peptide concentration was run in triplicate. See Table 1.5 for further details.

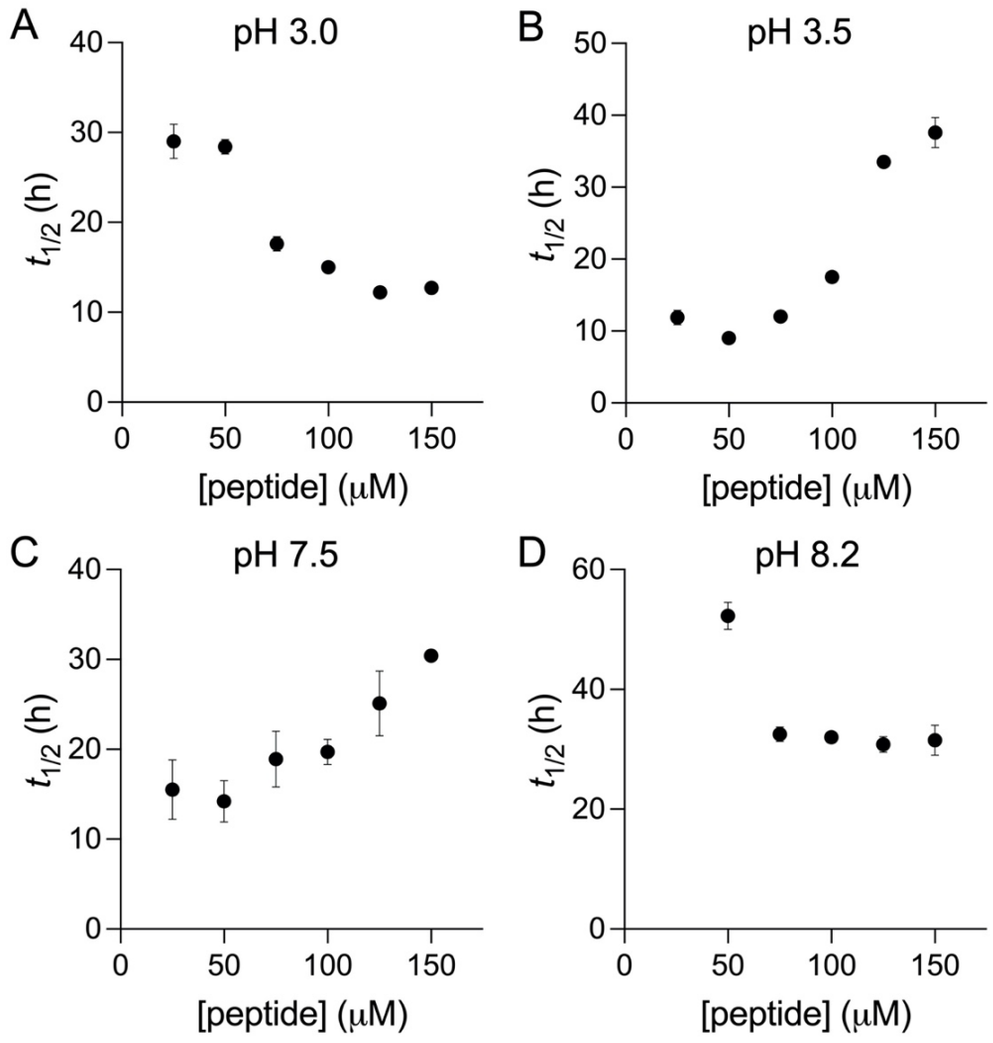


Figure 1.19. Dependence of the GLP-1 aggregation half-time on the peptide concentration.

Data collected by Dr Karolina Zapadka. Dependence of the half-time on the concentration of GLP-1 in 25 mM citrate buffer at pH 3.0 (A), 25 mM citrate buffer at pH 3.5 (B), 25 mM phosphate buffer at pH 7.5 (C), and 25 mM Tris buffer at pH 8.2 (D). All samples incubated at 37 °C. The error bars shown in A-D correspond to the standard deviations of the kinetic parameters determined from one experiment where each peptide concentration was run in triplicate. See Table 1.5 for further details.

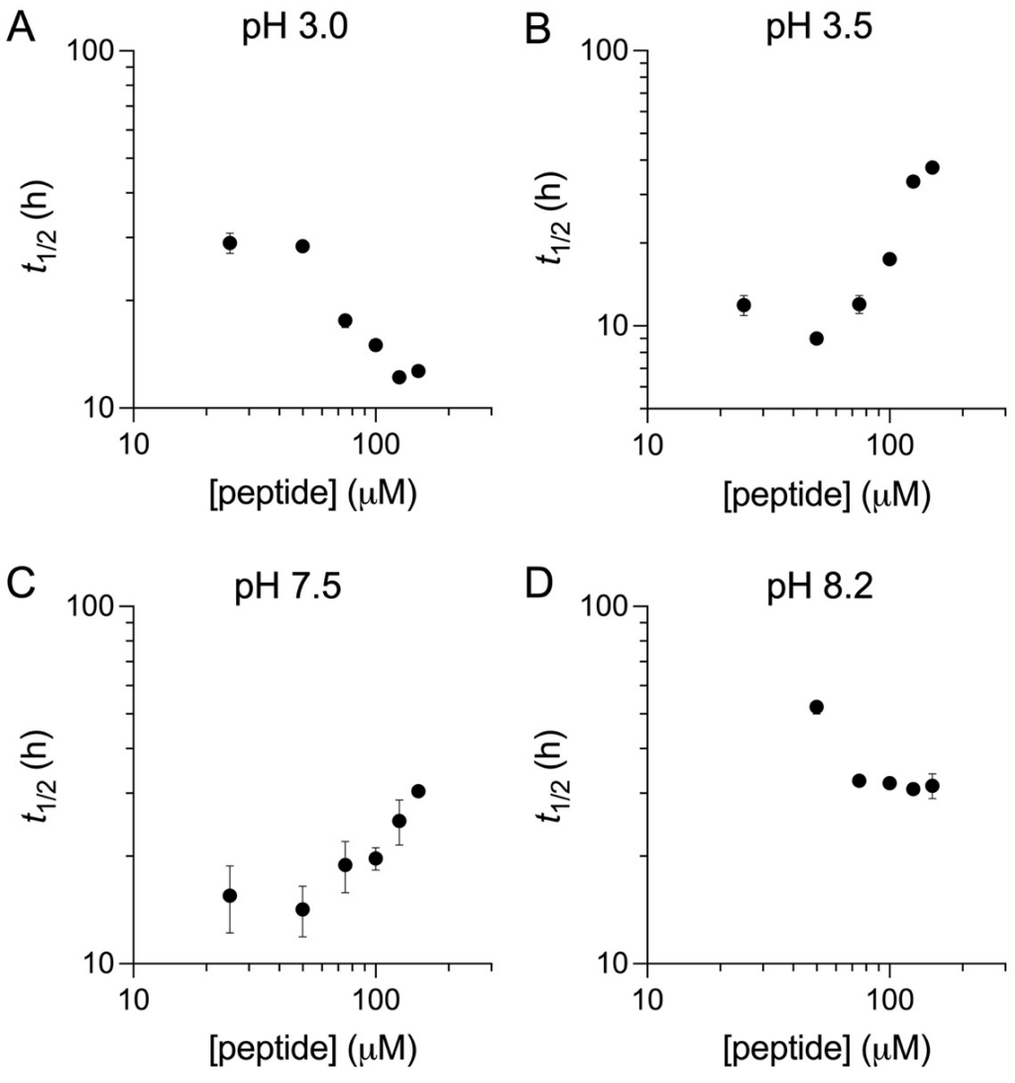


Figure 1.20. Dependence of the GLP-1 aggregation half-time on the peptide concentration on a log-log scale.

Data collected by Dr Karolina Zapadka. Dependence of the half-time on the concentration of GLP-1 in log-log scale in 25 mM citrate buffer at pH 3.0 (A), 25 mM citrate buffer at pH 3.5 (B), 25 mM phosphate buffer at pH 7.5 (C), and 25 mM Tris buffer at pH 8.2 (D). All samples incubated at 37 °C. The error bars shown in A-D correspond to the standard deviations of the kinetic parameters determined from one experiment where each peptide concentration was run in triplicate. See Table 1.5 for further details.

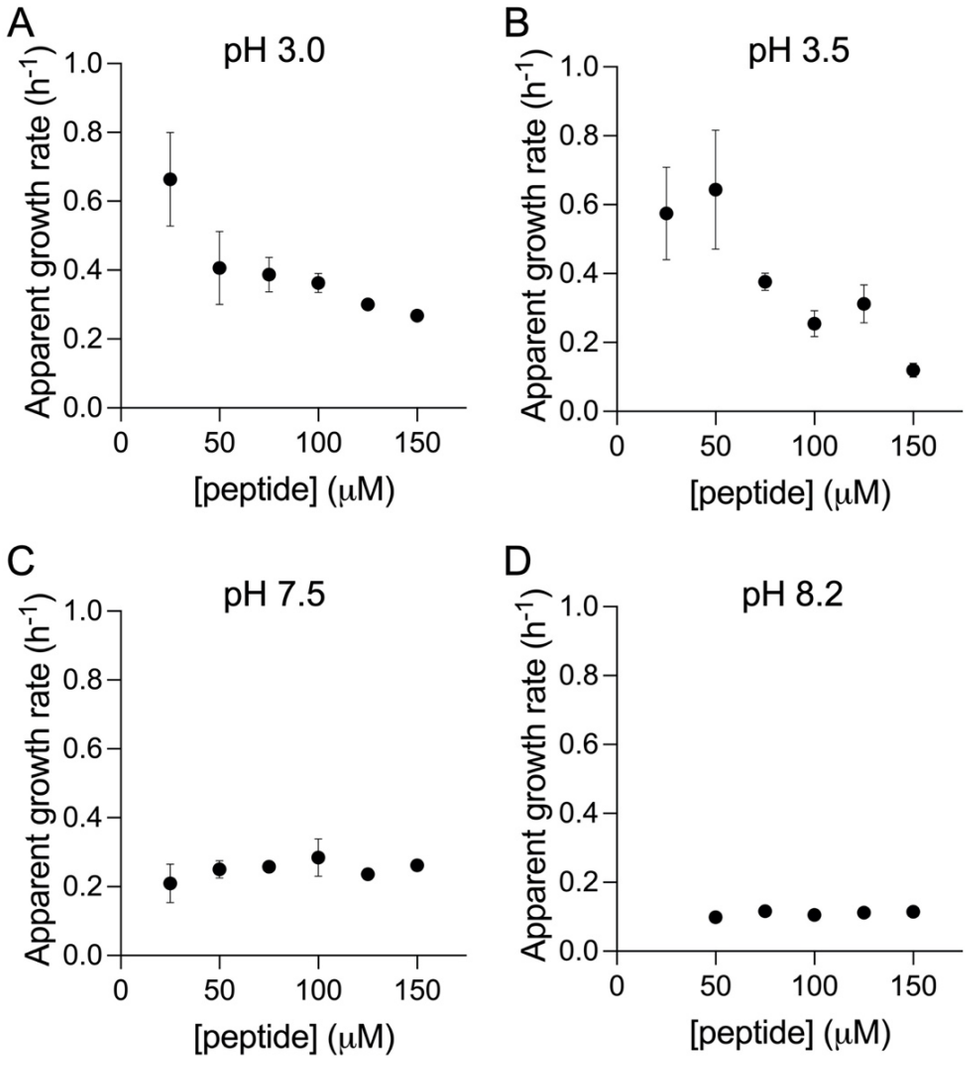


Figure 1.21. Apparent growth rate of GLP-1 as a function of peptide concentration at different pHs.

Dependence of the apparent growth rate on the concentration of GLP-1 in 25 mM citrate buffer at pH 3.0 (A), 25 mM citrate buffer at pH 3.5 (B), 25 mM phosphate buffer at pH 7.5 (C), and 25 mM Tris buffer at pH 8.2 (D). All samples incubated at 37 °C. The error bars shown in A-D correspond to the standard deviations of the kinetic parameters determined from one experiment where each peptide concentration was run in triplicate. See Table 1.5 for further details.

Petide concentration (μM)	Half-time (h)	SD	Lag time (h)	SD	Apparent growth rate (h^{-1})	SD
pH 3.0						
150	12.7	0.2	5.3	0.3	0.27	0.02
125	12.2	0.2	5.6	0.3	0.30	0.01
100	15.0	0.4	9.4	0.8	0.36	0.03
75	17.6	0.8	12.4	1.4	0.39	0.05
50	28.4	0.8	23.3	0.5	0.41	0.11
25	29.0	1.9	25.9	2.6	0.67	0.14
pH 3.5						
150	37.6	2.1	31.0	2.7	0.31	0.03
125	33.5	0.6	27.0	0.9	0.31	0.06
100	17.5	0.5	10.9	1.3	0.31	0.05
75	12.0	0.9	6.5	0.1	0.37	0.05
50	9.0	0.2	5.0	0.8	0.53	0.10
25	11.9	1.0	9.0	0.5	0.72	0.15
pH 7.5						
150	30.4	0.3	22.7	0.4	0.26	0.02
125	25.1	3.6	16.6	3.6	0.24	0.00
100	19.7	1.4	12.5	0.2	0.28	0.06
75	18.9	3.1	11.1	3.2	0.26	0.02
50	14.2	2.3	6.1	2.5	0.25	0.03
25	15.5	3.3	5.6	1.8	0.21	0.06
pH 8.2						
150	31.5	2.5	14.6	0.3	0.12	0.02
125	30.8	1.3	12.9	0.6	0.11	0.01
100	32.0	0.7	13.0	0.8	0.11	0.01
75	32.5	1.2	15.3	0.3	0.17	0.01
50	52.3	2.3	31.9	1.6	0.10	0.00
25	/	/	/	/	/	/

Table 1.5. Summary of the averages and standard deviations of the kinetic parameters obtained from multiple repeats of the ThT assays for GLP-1 at different pH values.

Data collected by Dr Karolina Zapadka.

1.11 Changes in Secondary Structure During Aggregation

To study the change of the secondary structure during the aggregation process, time-course measurements of GLP-1 were conducted in the Jackson Group. The measurements of far-UV CD spectra at different time points from 0 up to 144 h show for all four pH conditions (pH 3.0, 3.5, 7.5 and 8.2) an overall change from a structure containing α-helical, turn, disordered and β-sheet structure to a structure containing mainly β-sheet structure.

In citrate buffer at pH 3.0, the most dominant change of the secondary structure can be observed between 4-6 h, and therefore approx. the same time as the lag time of 5 h, Figure 1.22A & Table 1.5. The time-course measurements at pH 3.5 show an earlier change in the secondary structure. Here, the main change is observed after less than half of the measured lag time of 31 h, Figure 1.22B and Table 1.5.

Under neutral pH conditions (pH 7.5), the data shows again a close relation between the lag time and the significant change of the secondary structure, which both were determined to be around 23 h, Figure 1.22C & Table 1.5. However, a more gradual transformation in secondary structure can be seen at pH 8.2, with some significant changes happening between 7 and 8 h and a complete transformation of the secondary structure already after 9 h, Figure 1.22D. Between 9 and 144 h, the secondary structure stays constant. The change in secondary structure occurs earlier in comparison to the measured lag time of the ThT assay, which was measured to be 15 h, Table 1.5.

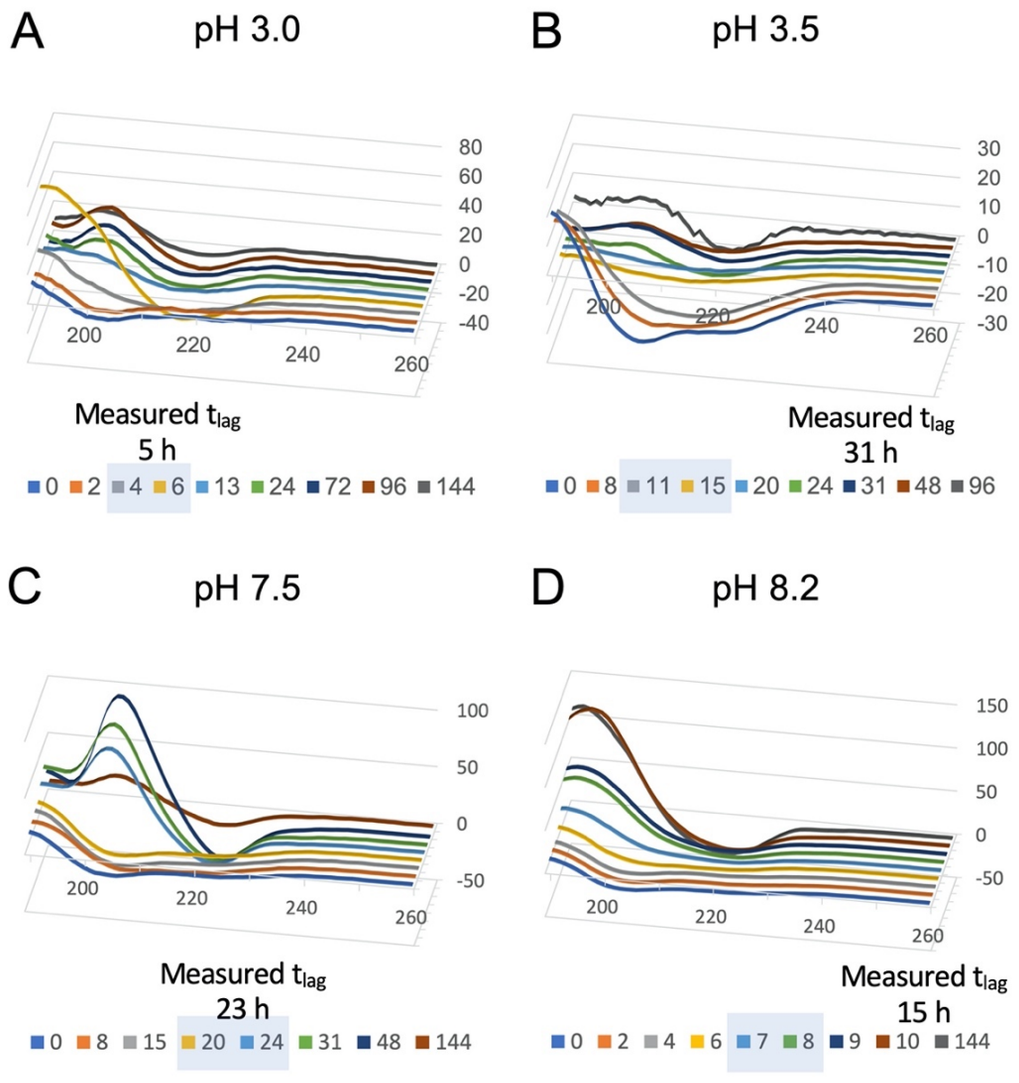


Figure 1.22. Changes in secondary structure in GLP-1 during aggregation shown with far-UV CD.

Measurements conducted by Dr Karolina Zapadka. Far-UV CD spectra of 150 μM GLP-1 incubated at 37 °C measured at different time points. The far-UV CD measurements were conducted at 25 °C. The x-axis and y-axis show the wavelength (nm) and the molar ellipticity (10^3 · deg · cm^2 · $dmol^{-1}$), respectively. The different colours represent different time points in hours. The measured lag times taken from previous ThT assays, Figure 1.17, are shown in black. Time points showing the first major changes of the secondary structure are shaded in blue. A) 25 mM phosphate, pH 3.0. B) 25 mM phosphate, pH 3.5. C) 25 mM phosphate, pH 7.5. D) 25 mM Tris, pH 8.2.

1.12 Thioflavin T as a Fluorescent Marker of Amyloid Fibrils

The study of amyloid fibrils was, from the beginning, influenced and inspired by the usage of dyes. The most widely used amyloid dye is the fluorescent dye thioflavin T (ThT). In 1959, its potential was demonstrated by Vassar and Culling.[281] Three decades later, the first detailed characterisation of the ThT fluorescence spectra and binding properties was undertaken by Naiki *et al*.[282] Their work on the fluorometric determination of amyloid fibrils *in vitro* showed that ThT in the absence of amyloid fibrils has only a faint fluorescence at the excitation and emission maxima of 350 and 438 nm, respectively. In comparison, the ThT fluorescence in the presence of amyloid fibrils was significantly increased at the shifted excitation and emission maxima of 450 and 482 nm, respectively. Understanding the specificity of ThT binding to amyloid fibrils has been studied extensively using a broad range of biophysical and biochemical analyses, protein/peptide engineering, fluorescence microscopy, and computational simulations.[283]

It has been suggested that ThT behaves in solution as a "molecular rotor" because of the single shared carbon-carbon bond between the benzylamine and benzathiole ring, Figure 1.23A.[284,285] If the rotation is possible, it enables the molecule to undergo a charge transfer which leads to the population of a non-fluorescent twisted intramolecular charge transfer state (TICT state)[284]. This quenching leads to a low fluorescence emission for non-immobilised ThT molecules. The binding of the ThT molecule to a fibril and selective immobilisation of a subset of the ThT molecules suppresses the population of the TICT state, which leads to an increase in the ThT fluorescence in samples with amyloid fibrils compared to samples without.

The characteristics of increased fluorescence caused by the binding of ThT to amyloid fibrils has brought the mechanism of its binding into the focus of a wide range of studies. The typical structure showing the fibril-ThT interaction has been described by a "channel" model.[286–288] Channel-like motifs are formed by amino-acid side chains forming cross-strand ladders. In these channels, the

ThT molecules bind aligned parallel to the long axis of the fibre, Figure 1.23B.[289] It was shown that ThT molecules prefer certain types of channels, especially channels formed by hydrophobic or aromatic residues.[290] In contrast, highly charged fibrils often show a poor affinity for ThT.[291] Furthermore, the minimum length of a channel to bind a ThT molecule was determined experimentally and corresponds to the height of five or more β-strands,[292] whereas computational simulations suggested that, in the absence of repulsive forces ,already a groove formed by the side chains of four β-strands is sufficient to bind ThT.[293]

It is assumed that the groove's minimum length is why ThT typically does not bind to β-sheet rich globular proteins as they usually contain highly twisted β-sheet motifs consisting of four or less β-strands.[289,294,295] However, from several significant studies on the specificity of ThT, it is known that there are some exceptions where ThT has been shown to bind to globular proteins. For example, it was shown that ThT binds to acetylcholinesterase.[296] In contrast to fibrils, here, the binding site is primarily formed of α-helices. The interaction of ThT and the hydrophobic binding pockets is again mediated by aromatic side chains, including π-stacking of Tyr and Trp.[296] Similarly, ThT has been shown to bind to a hydrophobic pocket of human serum albumin with an observed binding constant comparable to that of other drug-like molecules.[297]

Despite some binding to globular proteins, ThT binds relatively specifically to amyloid fibrils and shows low or no binding affinity to soluble proteins in folded, unfolded or partially folded states.[298]

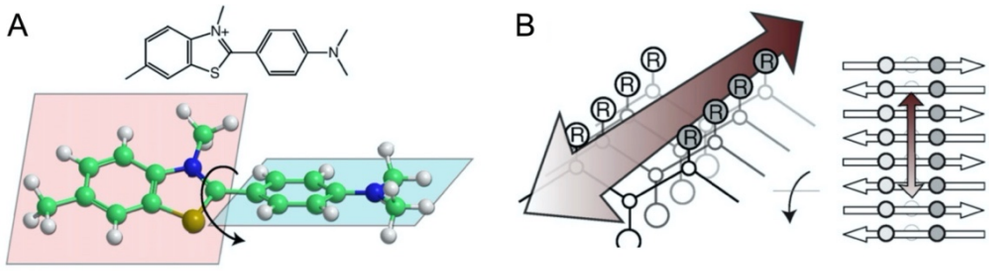


Figure 1.23. Structure of ThT and its binding model to fibril-like β-sheets.

A) Structure of ThT consisting out of two planar segments whose mutual rotation defines its chirality and influences its fluorescent properties. B) The binding of ThT molecules is proposed to occur along surface side-chain grooves. These so called “channels“ run parallel to the long axis of the β-sheet. Figure adapted from M. Biancalana & S. Koide, 2010.[283]

However, the low or sometimes complete lack of binding to non-fibrillar aggregates does not mean that these species cannot influence the ThT fluorescence indirectly. For example, it cannot be excluded that these aggregates reduce the overall fibril yield or affect the kinetics of fibril formation. In some cases, oligomers have been shown to have only a poor interaction with ThT resulting in a low ThT fluorescence.[299]

During ThT assays, it is always important to check if the ThT fluorescence is linearly related to the amount of fibrils in solution, i.e., peptide concentration. Another factor influencing ThT fluorescence is pH. Caution should be exercised when comparing measurements under acidic conditions and neutral pH values, as the protonation of the tertiary amine leads to a lower ThT fluorescence absorption and emission below pH 3.0.[300] Another critical factor influenced by pH is hydroxylation. Most notable under elevated temperatures and basic conditions, it can also play a role at neutral pH and 37 °C. For example, at pH 7.3 and 15 h incubation at 50 °C, a ~25% decrease in the ThT absorbance has been attributed to hydroxylation.[301]

It has also been suggested that ThT binds to at least two different binding sites in insulin amyloid fibrils which each have a different influence on its fluorescence,[302] which is a good reminder that a change in ThT fluorescence intensity in a fibrillar sample does not automatically mean a change in the total fibril concentration. Instead, it could simply be caused by an equilibration process between two ThT binding sites on a given fibril surface.

ThT assays can be conducted *in-situ* or *ex-situ* as fixed-timepoint assays. The latter has the advantage that no ThT is present during the actual aggregation process as it can theoretically influence the process. However, during fixed-timepoint assays, the extraction of samples, volume change, and agitation caused by sample extraction can also influence the aggregation process. A recent study showed that ThT present during the aggregation process alters the fibrillation kinetics of α-synuclein in comparison to an *ex-situ* assay.[303] However, the main differences were only seen for the fluorescence intensity. The primary kinetic parameters lag time and $t_{1/2}$ were not significantly affected. One possible explanation for the difference in the fluorescence intensity is the higher accessibility of binding sites during an *in-situ* assay.[304]

Overall, *in-situ* ThT assays are a powerful tool to follow the kinetics of fibril formation. They enable detailed kinetic analyses of the aggregation process, however, the linear dependence of the ThT fluorescence on fibril concentration should be checked and additional factors such as the effect of ThT binding to excipients and globular proteins (if included in the assay) must be excluded. It is also always recommended to support conclusions from ThT assay data with orthogonal methods such as AFM, SEM, TEM, far-UV CD or FT-IR.

1.13 Measurement of the Secondary Structure of Peptides Using Circular Dichroism

Biomolecules often possess a molecular asymmetry which means that their mirror image is not identical. Such behaviour is known as chirality, and these molecules are said to be optically active. Circular dichroism exploits the fact that chiral molecules have a non-zero difference in response to left- and right-circular polarized light.[305] In the case of polypeptides, the main chromophores interacting in the far-UV region are the peptide bonds.[306] If the latter is part of a regular folded structure as in an α-helix or β-sheet, it is influenced by the conformationally asymmetric structure.

To measure the circular dichroism, an unpolarised light source gets polarised and subsequently circular polarised with a quarter-wave modulator.[305,307] Depending on the modulation the circularly polarised radiation can be comprised of one of two different circularly polarised vectors.[305] Both vectors, the right-handed and left-handed, are of equal intensity. Samples with chromophores in optically symmetrical structures will typically absorb the two components in equal measures. In contrast, chromophores in an asymmetric environment will absorb the two different vectors to a different extent. The difference between the absorption of the left-handed and right-handed polarised light provides insights into the conformation of the measured samples. The method of far-UV CD profits from requiring only a low sample volume needed to gain structural data and the fact that the samples can be measured in solution. Figure 1.24 shows different components of typical far-UV spectra of a polypeptide adopting different conformations. Excellent reviews or book

articles of the CD technique and its applications to the study of proteins and other biomolecules have been written by L. A. Nafie (1995)[305], Kelly *et al*. (2005)[307] and S. R Martin & M. J. Schilstra (2008).[308]

A standard method to analyse far-UV CD spectra is to compare the measured dataset with spectra of proteins whose crystal structures have been determined. This is accomplished with the software package CDPro and the web-based Dichroweb, there are two prominent compilations of algorithms and reference sets.[309]

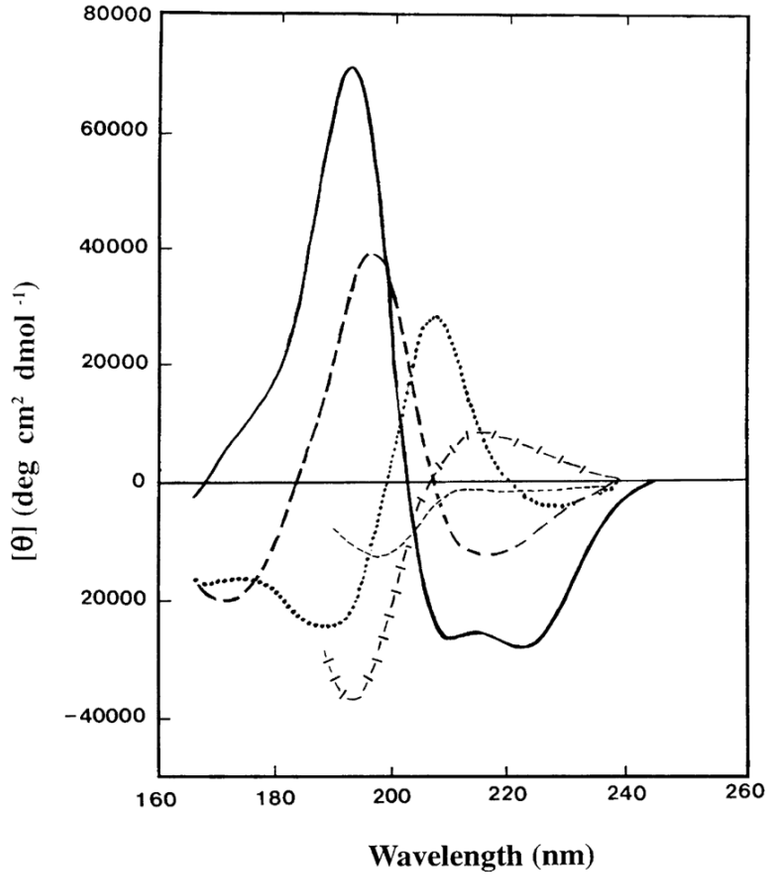


Figure 1.24. Far-UV CD spectra associated with different types of secondary structure of proteins.

Spectra shown of α-helix (solid line), anti-parallel β-sheet (long dashed line), type I β-turn (dotted line), extended 3_1-helix or poly (Pro) II helix (cross dashed line), irregular structure (short dashed line). Figure adapted from Kelly *et al*., 2005.[307]

1.14 Fourier-transform Infrared Spectroscopy of Polypeptides

A different spectroscopic approach to gain information about the secondary structure of polypeptides is Fourier-transform infrared (FT-IR) spectroscopy. While CD uses circularly polarised light, FT-IR measures the wavelength and intensity of a sample's absorption of infrared radiation.[310] The technique is thereby based on interferometry and Fourier transformation.[311] The interferometer in the instrument has one fixed and one movable mirror. A beam splitter directs 50% of the light onto the static and 50 % onto the movable mirror. If the right path length difference is introduced, it results in interference as soon as the two beams are recombined. After the beam interacts with the sample, the measured light intensity can be set in relation to the position of the movable mirror.[311] It results in an interferogram, which is the Fourier transform of the spectrum. Therefore, a second Fourier transform is performed by a computer to convert the data back to a spectrum.[312] In contrast to other spectrophotometric approaches, for example, far-UV CD, FT-IR is not based on a monochromatic scan of the sample. FT-IR benefits from its polychromatic approach in terms of speed and sensitivity.[311]

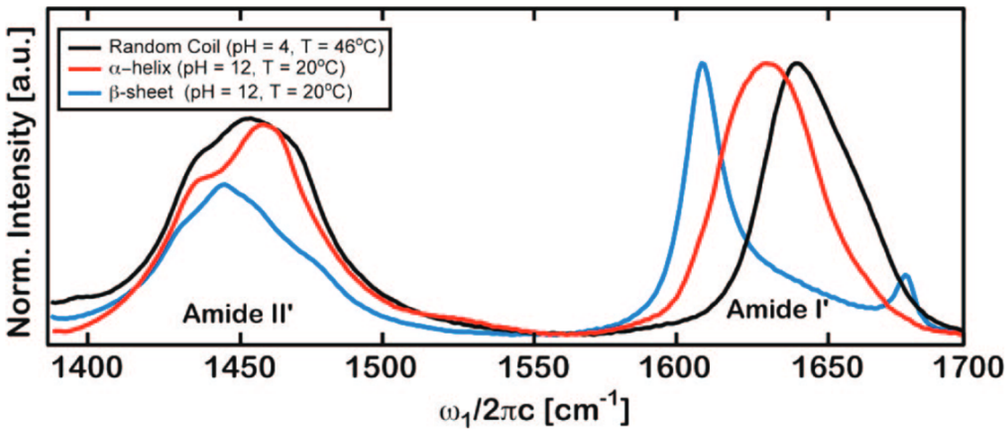


Figure 1.25. FT-IR spectra of poly-L-lysine as a function of pH and temperature.

The homopolymer forms a β-sheet at pH 12 and 46 °C, α-helix at pH 12 and 20 °C and random coil at pH 4 and 20 °C. Figure adapted from L. P. DeFlores *et al*., 2009.[313]

In the infrared spectroscopy of polypeptides, the peptide bonds absorb mainly in nine different characteristic infrared active vibration modes. The two most prominent vibration modes for analysing the secondary structure of polypeptides are the amide I and amide II modes.[312] At a frequency of approx. 1690-1600 cm^{-1}, the amide I mode absorbs mainly due to the C=O stretching vibrations. Additionally, minor contributions come from out-of-phase C-N stretching and the N-H inplane-bend. The amide II mode, in comparison, absorbs at a frequency of approx. 1575-1480 cm^{-1}. Its absorption is mainly caused by N-H plane-bend and the C-N stretching vibration.[312]

The modes shift depending on the structure of the sample. From previous studies, it is known that the amide I mode is composed from different absorption bands. A strong absorption band at 1610-1640 cm^{-1} and a weaker band at 1680-1690 cm^{-1} are typical for β-sheets. The absorption band for α-helical and random coil structure have been determined to be at 1640-1650 and 1650-1660 cm^{-1}, respectively. In Figure 1.25, typical FT-IR spectra of poly-L-lysine in the three different conformations, α-helical, β-sheet and random coil are shown.

1.15 Principal Component Analysis of Spectroscopic Data

The principal component analysis (PCA) is a computational method to analyse a set of multi-dimensional data,[314] for example, the time-dependent change of absorbance spectra. The analysis aims to extract the important information from data by finding inter-correlated quantitative dependent variables. The latter are orthogonal variables and called principal components (PCs).[315] In the case of a matrix of time-dependent absorption spectra, a PCA can give an overview of the variables which contribute the most to the different characteristics of the spectra at different time points.

A mathematical approach to deconvolute a data table in which each row describes an absorbance spectrum is singular value decomposition. After such a deconvolution, the data can be transformed back to the original data space,[316] such as the absorbance spectra space. It can help to understand the dominant

spectra composing the measured data. For example, the contribution of the first and most dominant principal component to each spectrum can be calculated using only the eigenvalue of the first principal component while all other eigenvalues are set to zero. The resulting matrix, M_{PC1} contains then, for each time point, the contribution of the first principal component, Figure 1.26.

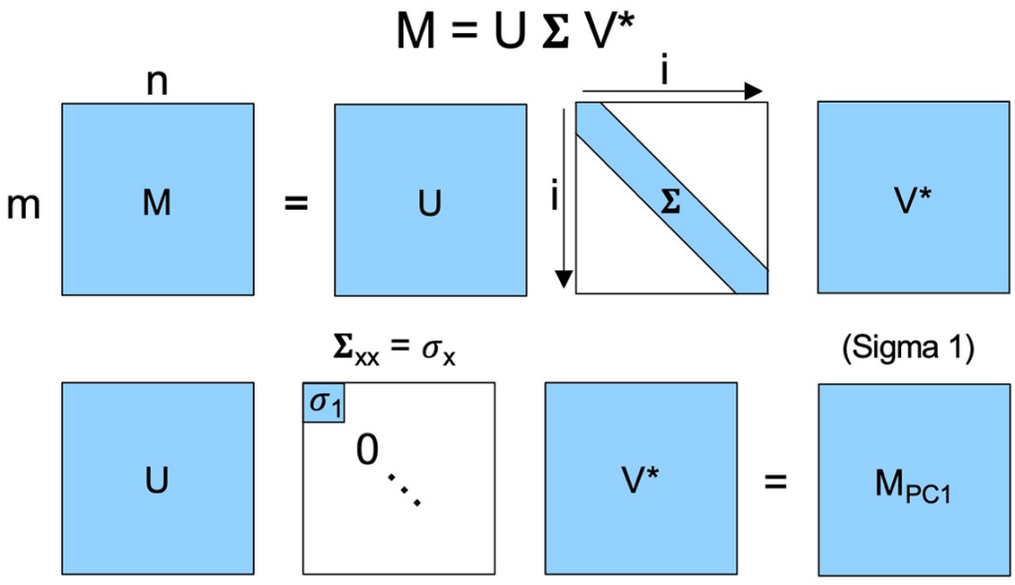


Figure 1.26. Singular value decomposition and back transformation of the first principal component to the original data space.

Decomposition of a real matrix M into a real unitary matrix U, a diagonal matrix **Σ** with the eigenvalues σ_x of the data set and a unitary matrix V. Multiplication of the two unitary with a single eigenvalue, for example σ_1, leads to the contribution of the first principal component for each row summarised in matrix MPC1 (Sigma 1).

1.16 Measurement of the pI of Polypeptides

The isoelectric point (pI) is the pH value at which the molecule carries no net charge. Therefore, the pI value is of particular interest in the context of peptide aggregation and was determined for all investigated peptides and lipidated peptides using imaged capillary isoelectric focusing (iCIEF).

iCIEF is an enhanced version of capillary isoelectric focusing (CIEF).[317,318] Both share the approach to focus polypeptides isoelectrically in a capillary coated with fused silica (FS). The coating is needed to suppress the electro-

endosomatic flow (EOF). This is essential because EOF disturbs the focusing process.[319]

At the beginning of the measurement, the capillary gets filled with a mixture containing carrier ampholytes covering the anticipated pH range (e.g., 3-8), the sample(s) and pI markers.[320] Sometimes also blocking agents are included, which have a lower or higher pI compared to the samples measured. Typical examples are iminodiacetic acid (pI 2.2) and arginine (pI 10.7).[319] When the pI and the amount of the blocker agents are chosen correctly, this can help focus the mixture correctly. The pH of the starting mixture is determined by the concentration and dissociation state of all substances present. The capillary is connected to an inlet and outlet vial filled with a low pH solution (H_3PO_4) and a high pH solution (NaOH). A positive voltage can be applied to the inlet, which becomes the anode. The outlet is grounded and acts, therefore, as the cathode.[320]

As soon as the voltage is applied, highly mobile hydrogen ions move towards the cathode, and hydroxyl ions move in the opposite direction. With time a pH gradient is created, and amphoteric molecules protonate/deprotonate and focus in a small zone in which their net charge is minimised.[319] To determine if the molecules are in focus, CIEF relies on the current as an indirect measurement. iCIEF profits in this case from the fact that focusing is directly detected. Another advantage of iCIEF compared to CIEF is that it needs no mobilisation step, which increases the accuracy of the measurement.[317,319]

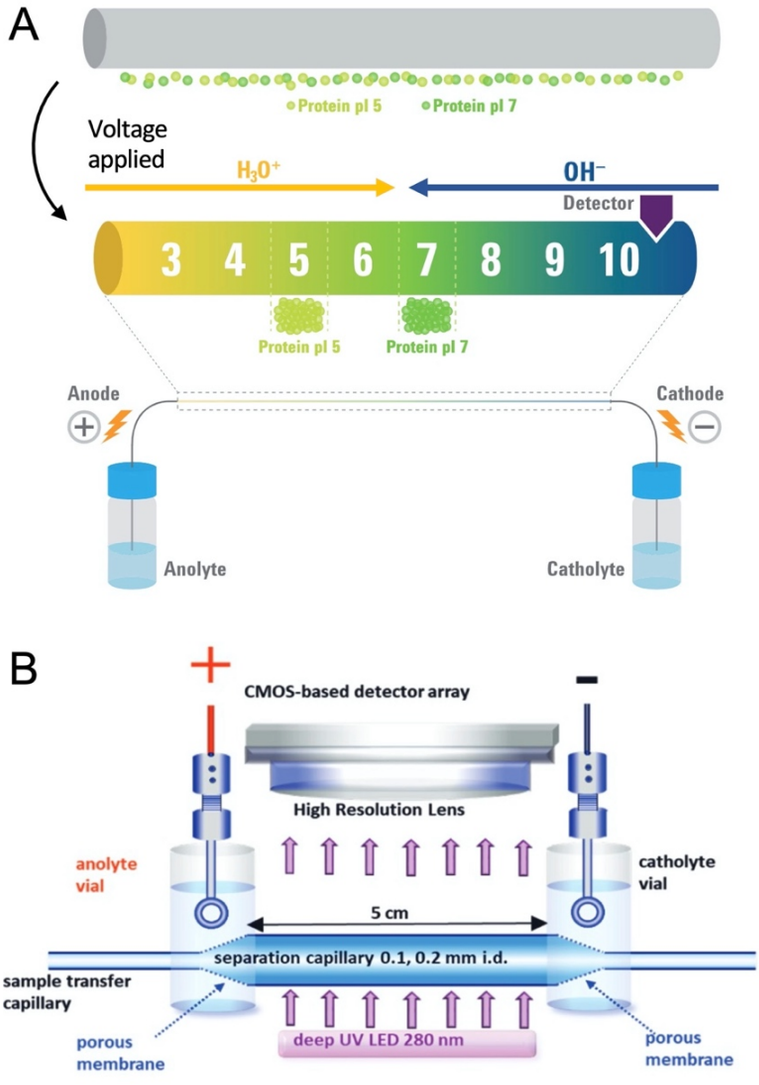


Figure 1.27. Schematic of capillary isoelectric focusing (CIEF) and imaged CIEF (iCIEF).

A) Depiction of the capillary composition before and after the focussing process during a CIEF run. Before the voltage is applied, the mixture of carrier ampholytes, blocking agents, the protein or peptide sample to be analysed, and pI markers are all homogeneously distributed over the whole capillary. The applied voltage leads with time to a pH gradient within the capillary. In a classical CIEF experiment, the neutralised and focussed molecules must be subsequently mobilised and moved past the point of detection, a spectrophotometric detector (UV-DAD) at 280 nm. Figure adapted from T. Kristl, H. Stutz, C. Wenz & G. Rozing, 2014.[320] B) Schematic of iCIEF. Compared to CIEF, it enables the measurement of the UV signal over the whole range of the capillary. It leads to an improved focus because the focusing process can be observed not only based on the current drop to a minimum as for CIEF but in real-time during the whole process. The precision also gains from the fact that no mobilisation step is needed to move the sample past a POD. Figure adapted from G. Rozing, 2019 (Picture courtesy of H. Stutz, Salzburg University).[319]

1.17 Aims of the Thesis

The aim of the work presented in this Thesis was to look into the aggregation pathways of GLP-1 and chemically modified forms of GLP-1 and to characterise their physical stability. GLP-1 and its analogues were selected due to their importance as therapeutics as well as their size, chemical structure and propensity to form fibrils. The focus was on the influence of amidation of the C-terminus as well as two lipidated versions which were investigated to understand how lipidation influences the aggregation behaviour of the peptide.

Objective 1

Based on the work of Dr Karolina Zapadka, the first objective of the research was to study the aggregation of GLP-1 at different time points during aggregation in order to understand the size, structure and nature of species in solution throughout the aggregation reaction. Motivated by the pH-dependent concentration dependence of the aggregation kinetics of GLP-1, four different pH values were selected which showed different behaviour. Many different biophysical techniques were used in order to obtain a better understanding of the aggregation mechanisms.

Objective 2

Another focus of the Thesis was on the determination of the influence of C-terminal amidation (Am-GLP-1) on the physical stability and aggregation kinetics of GLP-1. The effect of pH and peptide concentration was studied and compared in detail to the results obtained for GLP-1. The amidation of the C-terminus changes the net charge of the peptide enabling the role of net charge in determining the solubility as well as the aggregation propensity to be established.

Objective 3

Two lipidated forms of Am-GLP-1 were studied to determine the influence of lipidation on the aggregation kinetics. Similar to Am-GLP-1, the impact of pH and peptide concentration on the aggregation kinetics/amyloid fibril formation was studied. The studies also provided information on how the lipidation influences the secondary structure of the peptide by a comparison of results with those obtained for GLP-1 and Am-GLP-1.

Objective 4

Analysing the collected data using different fitting approaches provided a better understanding of the underlying mechanisms of the aggregation process. ThT assay data were subjected to a sigmoidal fit and also fit to an analytical solution of a Master equation describing the formation and growth of fibrils by nucleation, elongation and fragmentation.

2 METHODS AND MATERIALS

2.1 Peptides

The 31-residue GLP-1 (7-37) and Am-GLP-1 (7-37) (HAEGTFTSDVSSYLE-GQAAKEFIAWLVKGRG-OH and HAEGTFTSDVSSYLEGQAAKEFIAW-LVKGRG-NH_2) were purchased from Bachem (Purity: 98.5% and 96.7%). Both peptides have a molecular weight of approximately 3.36 kDa. IPP4 (HAEGTFTSDVSSYLEGK(γ-E-palmitoyl)-AAREFIAWLVRGRG-OH was purchased from Peptides International (Purity: 96.3%) and IPP5 HAEGTFT-SDVSSYLEGQAAK (γ-E-palmitoyl) EFIAWLVRGRG-NH_2) was purchased from Bachem (Purity: >95.0%). All peptides were used without further purification.

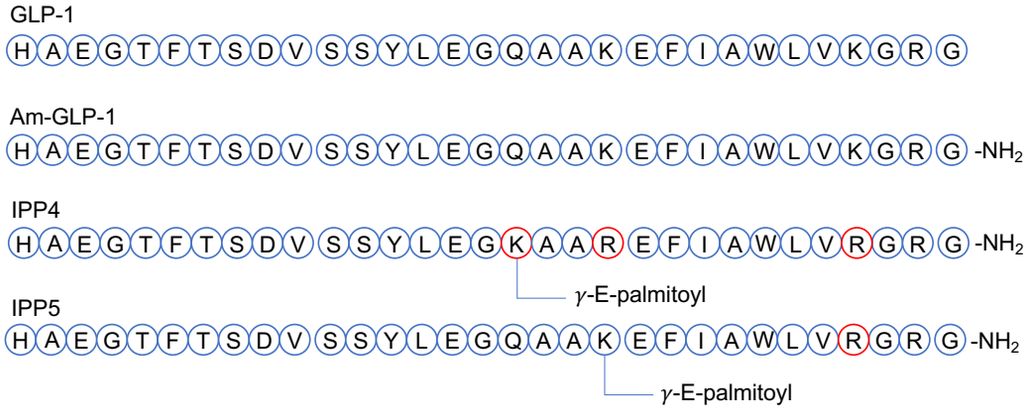


Figure 2.1. Amino-acid sequence of GLP-1 and three GLP-1 analogues.

The lipidated peptides are both based on the C-terminally amidated form of GLP-1 (Am-GLP-1) with additional modifications indicated by the red rings.

2.2 Buffers and other Reagents

The buffers were purchased from Sigma-Aldrich (citric acid anhydrous and sodium citrate dihydrate) and Baker (sodium phosphate dibasic 7-hydrate and dibasic sodium phosphate monohydrate/Tris base and Tris HCl). All buffers were prepared by weighing of the appropriate acidic and basic forms of the buffers, which were subsequently dissolved in deionised water. The final buffer concentration was 25 mM and citrate buffer was used at pH 3.0, 3.5, 4.0, 4.5, 5.0, 5.5 and 6.0, phosphate buffer at pH 6.5, 7.0 and 7.5 and Tris buffer at pH 8.0, 8.5 and 9.0. The pK_a of Tris buffer is temperature-dependent which means if the temperature is increased from 25 to 37 °C, the pH of the buffer changes from 8.0, 8.5 and 9.0 to 7.7, 8.2 and 8.7, respectively. In this Thesis, the samples in Tris buffer are referred to as samples at pH 7.7, 8.2 and 8.7 if not stated otherwise, as these are the pH values under aggregation conditions. No additional salt was added. The ionic strengths due to the buffer itself varied with pH and were in the range from 2 to 121 mM, Figure 2.2. The decision to conduct all measurements at the lowest possible ionic strength at each pH was based on a brief study by Dr Karolina Zapadka, which showed no significant differences in aggregation rate over this ionic strength range, Section 1.6.3.

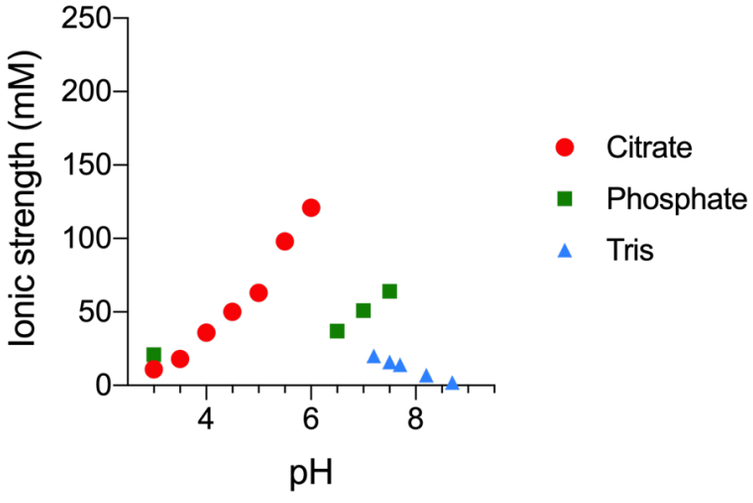


Figure 2.2. Ionic strength of citrate, phosphate and Tris buffers

The calculated ionic strength of 25 mM citrate, phosphate and Tris buffer prepared at 25 °C and used at 37 °C. The calculation was performed with the BioMol buffer calculator.[321]

After preparation, the buffers were filtered through a 0.2 µm PES filter unit (Thermo Scientific™ Nalgene™ Rapid Flow™ 90 mm) and the pH checked using a multiparameter pH meter (SevenExcellence, Multiparameter, Mettler Toledo, pH electrode: InLab® Micro-Pro-ISM). Buffers were stored at -20 °C.

The sources of all other reagents are given in the relevant Methods sections.

2.3 UV-visible Absorption Spectroscopy – NanoDrop

A NanoDrop 2000 spectrophotometer (Thermo Scientific) with a 1 mm pathlength was used to determine the concentration of GLP-1 or ThT in solution. After recording an appropriate buffer as a blank, an absorption spectrum of the sample was recorded in triplicate. Concentrations were determined using the measured absorbance at 280 nm and corrected for light scattering using the absorbance at 340 nm.

$$c = \frac{A_{280} - A_{340}}{\epsilon_{280} \cdot l} \tag{2.1}$$

The correlation of the absorbance with the concentration c is given by the Beer-Lambert equation, Equation 2.2, where ϵ_λ is the extinction coefficient at a specific wavelength and l the pathlength in cm.

$$A_\lambda = \epsilon_\lambda \cdot l \cdot c \tag{2.2}$$

For all GLP-1 peptides a theoretical extinction coefficient of 6990 $M^{-1}cm^{-1}$ at 280 nm was used, which was calculated using the Expasy ProtParam tool[322]. The molar extinction coefficient for ThT was 31600 M^{-1} cm^{-1} at 412 nm.[323]

The sample size was 2.5 µL. The data were analysed with the program NanoDrop 2000/2000c, which calculates the absorbance using the logarithmic relation between the absorbance A at a certain wavelength, and the intensity of the sample I_s and blank I_b

$$A_\lambda = -log\left[\frac{I_s}{I_b}\right] \tag{2.3}$$

2.4 Sample Preparation

The peptides were dissolved in 25 mM citrate, phosphate or Tris buffer up to a final stock concentration of 300 µM depending on the solubility of the peptide. Citrate was used to buffer from pH 3.0 to 6.0, phosphate from pH 6.5 to 7.5 and Tris from 8.0 to 9.0. After the peptides were dissolved, the pH was checked using a pH meter (SevenExcellence, Multiparameter, Mettler Toledo, pH electrode: InLab® Micro-Pro-ISM) and the solutions were subsequently filtered using a 0.45 µm PVDF Millipore centrifuge filter prior to use (Biofuge pico, Heraeus, 2 min, 13000 rpm at room temperature). The peptide concentration was then determined spectrophotometrically using a NanoDrop spectrophotometer, Section 2.3.

2.5 Solubility Measurements

The peptides were weighed, and approx. 0.2 mg were dissolved in 50 µL of buffer to get a final concentration of approx. 1200 µM. 25 mM buffer was used for all pH values between pH 3.0 and 9.0. After the buffer was added to the peptide, the samples were left to dissolve for 30 min and were subsequently centrifuged for 5 minutes at 13000 rpm at room temperature (Biofuge pico, Heraeus) before the concentration of the supernatant was determined spectrophotometrically using a NanoDrop spectrometer.

2.6 Liquid Chromatography-Mass Spectrometry (LC-MS)

For the LC-MS measurements, lyophilised sample was dissolved in 33% deionised water, 33% acetic acid (glacial, purum) and 33% acetonitrile (HPLC grade). The sample size was 5 µL with a peptide concentration of 30 µM. The LC-MS was performed using an Agilent System, with an Agilent Polaris analytical C8-A, 4.6 x 100 mm, 5 µm Agilent Technologies column attached. The experiment was run at room temperature. The instrument was equipped with a 3100 Mass Detector (Waters, GB) and the peptides were run according

to the methods summarised in Table 2.1 depending on whether they were lipidated or not.

Flow rate	1.5 ml/min	
UV detector	210 nm	
Run time	6 min	
Sample injection	5 µL	
Mobile phase *A*	0.1 % trifluoroacetic acid (TFA) in water	
Mobile phase *B*	0.1 % trifluoroacetic acid (TFA) in acetonitrile	
Gradient of the run for non-lipidated peptides		
Time (min)	Mobile phase *A* (%)	Mobile phase *B* (%)
0	90	10
5	10	90
5.75	10	90
6	90	10
Gradient of the run for lipidated peptides		
Time (min)	Mobile phase *A* (%)	Mobile phase *B* (%)
0	85	15
5	0	100
5.75	0	100
6	85	15

Table 2.1. Parameters for the LC-MS experiments.

2.7 Dynamic Light Scattering (DLS)

For the DLS experiments, a Zetasizer Nano ZS (Malvern Instruments, Malvern, UK) was used. Using a monochromatic laser light source and quasi-elastic light scattering it determines the mutual translational diffusion coefficient D_t of macromolecules in a sample. Assuming that all particles are spherical, homogeneous and with equivalent density (Mie scattering), the Zetasizer software calculates the intensity-weighted mean hydrodynamic distribution and of the number size distribution.

Samples of 150 µM GLP-1, freshly prepared or fibrillar, were measured. The

measurements were recorded at 25 °C in a Zen 2112 cuvette and with a scattering angle of 173 degrees. Before each measurement, the cuvette was rinsed three times with deionised water and subsequently with ethanol (95%). After the cuvette was dried using compressed air, it was filled with approximately 60 µL of the sample. The submenu measurement type was set on *peptide*, and in the submenu the appropriate dispersant was chosen. This was either 25 mM citrate, phosphate or Tris buffer. For each sample, three consecutive traces were recorded and averaged.

2.8 Imaged Capillary Isoelectric Focusing (iCIEF)

The pI of the peptides was determined using iCIEF (proteinsimple - Maurice). All peptides were dissolved in Tris buffer at pH 8.5 with a final peptide concentration of 336 µM (approx. 1 mg mL^{-1}). For each run, 40 µL of the sample was diluted with 160 µL of Master Mix, which contained distilled water, methyl cellulose, broad-range ampholytes, arginine and different pI markers. The composition of the Master Mix varied for different peptide samples.

For two runs of Am-GLP-1, the Master Mix was mixed using 178 µL distilled water, 168 µL 1% methyl cellulose (proteinsimple), 19.2 µL pharmalytes® 3-10 (Sigma), 9.6 µL 500 mM arginine (prepared from L-Arginine monohydrochloride, Merck), 4.8 µL pI marker I (Trp-Asp-Asp-Arg, pI 4.05), and 4.8 µL pI marker II (Trp-Tyr-Lys-Lys, pI 9.99).

The Master Mix for the GLP-1 sample contained urea to prevent precipitation and was prepared for approximately eight samples. For the Master Mix with urea, 0.96 g of urea was weighed and dissolved in 500 µL distilled water, vortexed and subsequently heated for 5 minutes at 80 °C (heat block). Immediately afterwards 400 µL 1% methyl cellulose (proteinsimple) was added and the sample vortexed again. After the sample cooled down to room temperature, 80 µL pharmalytes® 3-10 (Sigma), 80 µL 500 mM arginine (prepared from L-arginine monohydrochloride, Merck), 20 µL pI marker I (Trp-Asp-Asp-Arg, pI 4.05), and 20 µL pI marker II (Trp-Tyr-Lys-Lys, pI 9.99) were added.

The Master Mix for IPP4 and IPP5 followed the same method as GLP-1 but the pharmalytes® were changed to 40 µL pharmalytes® 3-10 (Sigma) and 40 µL pharmalytes® 5-8. Additionally, the second marker was changed to 20 µL pI marker II (pI 7.05).

The cartridge was prepared by adding 2 mL of Catholyte solution (100 mM sodium hydroxide in 0.1% MC, proteinsimple) to the OH^- electrolyte tank and 2 mL of Anolyte solution (80 mM phosphoric acid in 0.1% MC, proteinsimple) to the H^+ electrolyte tank. 2 mL 0.5% methyl cellulose (proteinsimple) and 500 µL fluorescence calibration standard (proteinsimple) were used as batch reagents.

All samples were measured with a pre-focusing time of 1 min at 1500 V and a focusing time of 10 min at 3000 V. Before the measurement a System Suitability (proteinsimple) sample was run to the check the performance of the instrument. The experimental data were processed and analysed using Compass Software for ICE (Protein Simple, USA).

2.9 Thioflavin T Aggregation Assay

Thioflavin T (ThT) was purchased from Sigma. For kinetic experiments, a stock solution of 300 µM of each peptide was prepared, as described above. The samples were incubated in a 96-well half-area plate made of black polystyrene with a clear bottom and a non-binding surface (Corning 3881, USA). The appropriate volumes of peptide stock solution and buffer were mixed with 6 µL of 1 mM ThT stock in the wells in the plate such that the final peptide concentrations were 25, 50, 75, 100, and 150 µM and the ThT concentration was 50 µM and the final volume 120 µL. A sealing tape (Corning, Thermowell™ Sealing Tape for 96 Well Plates, 6570) was used to protect samples from evaporation. Fluorescence measurements were carried out using a Fluostar Optima Microplate Reader (BMG Labtech, Offenburg, Germany) at 37 °C. Fluorescence was measured using an excitation filter at 440 nm and emission filter at 480 nm. Bottom reading of the plate every 30 min with 5 min of shaking prior to each measurement was performed. Each cycle was executed with the orbital shaker at 350 rpm, 5 flashes per well and fluorescence

measurements were made at a gain of 800. All measurements were made in triplicate for each peptide concentration in a single 96-well plate and each experiment was repeated at least three times on different days with freshly prepared samples.

2.9.1 Preparation of ThT Stock Solution

For all ThT assays, a 5 mM ThT stock solution was prepared by dissolving dry ThT powder (Sigma-Aldrich, T3516-5G, Lot # MKCF7580, dye content 66 %) in deionized water. After the ThT was completely dissolved, the stock solution was filtered using a 0.22 µm PVDF syringe filter. The stock was stored at -20 °C. Before use, the stock was diluted to 1 mM and the concentration was verified spectrophotometrically using a NanoDrop spectrometer and a molar extinction co-efficient (ε) of 31600 $M^{-1}cm^{-1}$ at 412 nM.[323]

2.9.2 ThT Assay Data Analysis: Fitting to a Sigma Function

Each individual ThT dataset was fit to

$$y = y_0 + \frac{A}{1 + \exp\left(-k\left(t - t_{\frac{1}{2}}\right)\right)} + bt \tag{2.4}$$

with y_0 as the starting fluorescence, A is the amplitude of the transition, $t_{1/2}$ the half-time, which is defined as the time at which the ThT fluorescence has reached 50% of its final baseline value, k the apparent growth rate and b the slope of the final baseline.[324,325] Note, in many other studies data are normalized prior to fitting such that the inclusion of the additional term bt is not needed. The lag time was calculated from the kinetic parameters obtained using:[324,325]

$$t_{lag} = t_{\frac{1}{2}} - \frac{2}{k} \tag{2.5}$$

2.9.3 ThT Assay Data Analysis: Fitting to a Fragmentation-Dominated Model

To gain further insight into the aggregation kinetics, the ThT data were also fitted to an equation based on principal moments. This simplified solution of a

Master equation takes primary nucleation, elongation and fragmentation into account.[83,280]

$$M_{RF}(t) = M_0\left[1 - \exp\left(C_- e^{-\kappa t} - C_+ e^{\kappa t} + \frac{\delta^2}{\kappa^2}\right)\right] \quad (2.6)$$

where the fibril mass is M_{RF}, the total protein concentration is M_0 and $C_\pm = \pm(\delta/2\kappa)$. Here δ and κ represent the effective rate for nucleated polymerisation without secondary pathways and the rate of multiplication of the filament population, respectively. The effective rate for nucleated polymerisation without secondary pathways, δ, is given by

$$\delta = \sqrt{2M_0^{n_c} k_n k_+} \quad (2.7)$$

It depends, therefore, on the total peptide concentration M_0, the size of the nucleus n_c, the primary nucleation rate constant k_n and the elongation rate constant k_p. δ is precisely the effective rate constant derived by Oosawa[81,326] for nucleated polymerisation without secondary pathways. The rate of multiplication of the filament population, κ, is given by

$$\kappa = \sqrt{2M_0 k_n k_m} \quad (2.8)$$

It includes the fragmentation rate constant k_m.

2.10 Far-UV Circular Dichroism (Far-UV CD) and CDPro

Circular dichroism measures the differential absorption of right-handed and left-handed polarised light by chiral molecules. The far-UV CD spectra of freshly prepared peptide samples (150 μM) in the appropriate buffer, as well as samples aggregated in a BMG Fluostar plate reader (150 μM) using the method described in Section 2.9 without ThT, were measured using a 0.1 mm pathlength cuvette in a Jasco J815 spectrophotometer (Jasco Spectroscopic Co. Ltd., Tokyo, Japan). The cuvette was filled with 75 μL of sample and the CD spectra and HT voltage were recorded from 180 to 260 nm at 25 °C. Data with a HT voltage higher than 500 V was discarded as the CD signal was too noisy to use. Therefore, only data from 190 to 260 nm were further analysed as in

this range the HT voltage was always below 500 V. For each measurement, spectra were acquired three times with a scanning speed of 20 nm min^{-1}.

After baseline subtraction, the spectra were processed using Spectra Manager Version 2 software and smoothed using a mean movement method with a convolution width of 5. The method, also known as the simple moving average method, uses a rectangular function as a weighting function. The smoothed value y(i) for each discrete value x(i) of the input signal was determined using Equation 2.9. The rectangle width is defined as N = 2m + 1 and i = m+1, m+2,…, n-m for n discrete values x(i).

$$y(i) = \frac{1}{N} \sum_{j=-m}^{m} x(i+j) \tag{2.9}$$

The data output was in ellipticity θ (mdeg). The mean molar ellipticity per residue $[\theta]_{MR}$ was calculated using Equation 2.10.

$$[\theta]_{MR} = 10 \cdot \frac{\theta}{(C_{MR} \cdot l)} \tag{2.10}$$

where l is the pathlength in cm, and C_{MR} is the mean residue concentration, which was calculated from the concentration of peptide, C, in molar units and the number of residues N in the peptide using Equation 2.11.

$$C_{MR} = C \cdot N \tag{2.11}$$

The data were further analysed using CDPro.[327,328] The programme compares a far-UV CD spectrum to the far-UV CD spectra of different proteins whose crystal structures and therefore secondary structures are known. Three different reference sets were used (SP29, SP43 and SDP48). The reference sets tested included data of 29, 43 or 48 proteins within the range of 178-260 nm (SP29) and 190-240 nm (SP43 and SDP48). A detailed comparison of the different reference sets has been made by Sreerama and Woody.[329] Their results indicate that far-UV CD data in the wavelength range of 190-240 nm can give reliable results when a larger reference set is used.[329] Whitmore and Wallace[309] give a complete list of the proteins in each reference set in their paper about the analysis of protein secondary structures and features of the DichroWeb, a web application of CDPro. The corresponding names of the datasets in DichroWeb are RDB1 (SP29), RDB4 (SP43) and RDB7 (SDP48). For GLP-1, the reference

dataset SP43 showed the most consistent results. Therefore, it was used to calculate secondary structure content using three different algorithms (SELCON3[327,330], CDSSTR[329,331,332] and CONTINLL[333,334]).

2.11 Fourier-transform Infrared Spectroscopy (FT-IR)

The Fourier-transform IR spectra of different peptide samples were measured with a Perkin Elmer FT-IR spectrometer (Spectrum 400) equipped with an Attenuated Total Reflectance (ATR) crystal. For each sample, small volumes (8 x 0.75 µL) were repeatedly dried on the crystal surface using a gentle, filtered air stream. The resolution employed was 4 cm^{-1} and the signal was accumulated using 32 repetitions. The spectra were recorded from 2000 to 1000 cm^{-1} and analysed in the amide I region (1720 - 1580 cm^{-1}) using the Perkin Elmer Spectrum Software.

2.12 Extrinsic and Intrinsic Protein Fluorescence

A Hitachi F-7000 fluorescence spectrophotometer plate reader was used for all extrinsic or intrinsic protein fluorescence time point measurements. Emission spectra were collected using top reading of the plate.

2.12.1 Tryptophan Fluorescence

The intrinsic fluorescence of 150 µM Am-GLP-1 at different pH values and at different time points was recorded. 100 µL aliquots were pipetted into wells of a 96-well half-area plate mode of black polystyrene, with flat wells and non-binding surface (Corning 3993, US). Emission spectra were collected from 285 to 500 nm using an excitation wavelength of 277 nm (scan speed: 1200 nm min^{-1}, excitation bandpass: 2.5 nm, emission bandpass: 2.5 nm, voltage: 900 V, response: 2.0 s, data point collection: 1 nm).

2.12.2 8-Anilinonaphthalene-1-sulfonic Acid (ANS) Binding Assay

8-anilinonaphthalene-1-sulfonic acid (ANS) has been used to detect the presence of exposed hydrophobic patches in the aggregated species.[335] ANS was purchased from Sigma-Aldrich (A1028-5G, Lot # MKCJ6296, purity ≥ 97 %). Measurements were taken at the beginning of the ThT assay and one after 144 h of aggregation. Typically, 50 µL of peptide sample was mixed with 2.5 µL 1 mM ANS stock directly in a 96-well half-area plate made of white polystyrene and non-binding surface (Corning 3693, US). The final concentrations were 143 µM (peptide) and 48 µM (ANS). The emission spectra were acquired from 420 to 600 nm using an excitation wavelength of 350 nm (scan speed: 240 nm min^{-1}, excitation bandpass: 1 nm, emission bandpass: 2.5 nm, voltage: 800 V, response: 0.01 s, data point collection: 1 nm).

Preparation of ANS Stock Solution

A 1 mM ANS stock solution was prepared by dissolving dry ANS powder in deionized water. After the ANS was completely dissolved, the stock solution was filtered using a 0.22 µm PVDF syringe filter. The stock was stored at -20 °C. The concentration was determined spectrophotometrically using a molar extinction coefficient at 350 nm of 5000 M^{-1} cm^{-1}.[335]

2.12.3 *Ex-situ* ThT Fluorescence Measurements

Ex-situ ThT measurements were taken before the aggregation assay and after 144 h of aggregation. Typically, 50 µL of peptide sample was mixed with 1 µL 2.5 mM ThT stock to a final concentration of 147 µM GLP-1 and 49 µM ThT in wells of a 96-well half-area plate made of black polystyrene, with flat wells and non-binding surface (Corning 3993, US). Emission spectra were collected from 450 to 600 nm using an excitation wavelength of 440 nm (scan speed: 240 nm min^{-1}, excitation bandpass: 5 nm, emission bandpass: 5 nm, voltage: 800 V, response: 0.01 s, data point collection: 1 nm).

2.13 Fibril separation

For the fibril separation experiment, 6 x 120 µL of GLP-1 or Am-GLP-1 at a concentration of 150 µM were prepared. The samples were aggregated at 37 °C with agitation for 144 h in a BMG Fluostar plate reader. After the incubation, the tryptophan fluorescence was measured using 3 x 100 µL of the aggregated peptide sample (Section 2.12.1). Subsequently, the samples were split in half, 2.5 µL of 1 mM ANS stock or 1 µL of 2.5 mM ThT stock was added to 50 µL of peptide sample, and the fluorescence spectra recorded as described in Sections 2.12.2 and 2.12.3.

In addition, 360 µL of aggregated peptide sample was placed into a 0.6 mL Eppendorf tube and centrifuged for 20 min at 17000 g and 25 °C. The supernatant was then carefully separated from the pellet using a pipette, Figure 2.3. The separated pellet was subsequently suspended in the appropriate buffer to reach a final volume of 360 µL. After a repetition of this process, the tryptophan, ANS and ThT fluorescence were measured as described in Section 2.12.2 and 2.12.3.

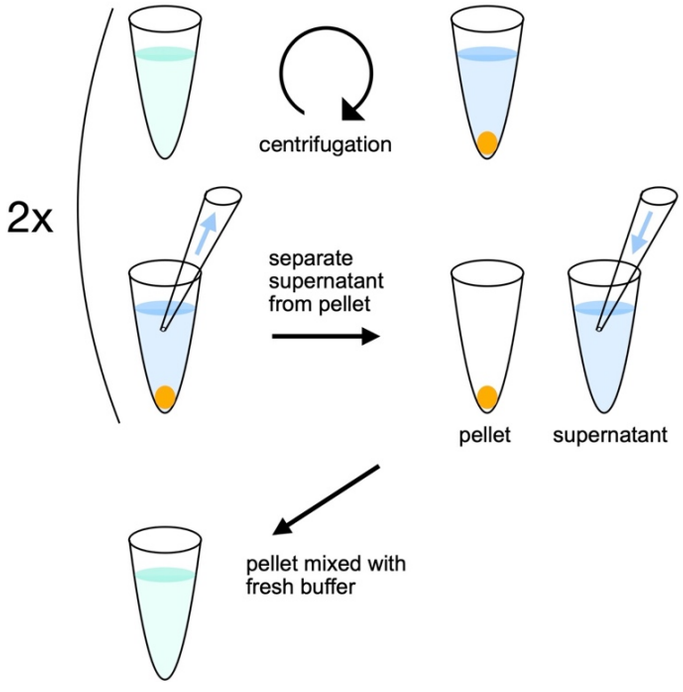


Figure 2.3. Separation of peptide aggregates from supernatant using centrifugation.

The separation of the supernatant was repeated twice.

2.14 Imaging

2.14.1 Atomic Force Microscopy (AFM)

Peptide samples in solution were taken directly after the peptides were dissolved and filtered and at different time points during the aggregation assay or at the end. Typically, 5 µL of each sample was spread onto freshly cleaved mica. After the sample dried, it was rinsed with 2 x 300 µL deionised water and dried using a gentle nitrogen stream. The AFM images were acquired with a PicoPlus AFM instrument with a PicoSPM II controller from Molecular Imaging (Agilent Technologies, USA) in AC mode (Tapping mode), equipped with a MikoMasch NSC26/No Al cantilever, between 65 and 130 Hz frequency, force constant varying from 0.6 to 2.0 N m^{-1} (Innovative Solutions Bulgaria Ltd., Sofia, Bulgaria). All images were taken in air and at room temperature. The images were analysed with the open-source software Gwyddion.[336] The stated width of the measured aggregates is equivalent to the full width half maximum (FWHM). Therefore, the profiles were fitted with a Gaussian and the standard deviation of the function multiplied by 2.35.[337,338]

2.14.2 Scanning Electron Microscopy (SEM)

SEM does not have the highest possible magnification of available microscopy techniques, but it has the advantage that it enables scanning over a broad range of magnifications and therefore it is less likely to miss important details in the sample. For example, if the sample spot has a diameter of 4 mm, for three 10 x 10 μm^2 pictures with AFM only a few millionths of the whole sample surface is imaged. As for every other microscopy technique using dried samples, the images are, of course, influenced by the drying process.

Typically, 5 µL of GLP-1 or Am-GLP-1 in buffer were spread onto microscope glass slides which were cleaned before use with acetone and isopropanol. The samples were either left to dry or the drop was gently blown away with a stream of nitrogen after it had been on the surface for 2 min. The slides were rinsed with 2 x 300 µL of deionized water and dried under a gentle stream of nitrogen before the samples were coated with a 10 nm thick platinum layer (Quorum

Technologies Q150T ES Turbo-Pumped Sputter/Carbon Coater). The samples were imaged using a Tescan Mira3 FEG-SEM. The In-Beam SE detector was used with SEM HV 5.0 kV and WD 4 m.

2.15 Net Charge and pI Determination

The net charge was determined using the Henderson-Hasselbalch equation and the isoelectric point calculation.[339] For residues of the amino acids with a negative charge the contribution can be calculated using Equation 2.12 and their individual pK_a values:

$$\sum_{i=1}^{n} \frac{-1}{1 + (10^{pK_a - pH})} \tag{2.12}$$

For positively charged residues the charge contribution can be calculated using the appropriate pK_b values as shown in Table 2.2 and Equation 2.13:

$$\sum_{i=1}^{n} \frac{1}{1 + (10^{pH - pK_b})} \tag{2.13}$$

The final charge is the sum over all negative charges plus the sum of all the positive charges which are, of course, both pH dependent.

Name	**1-Letter Symbol**	**pK_a**	**pK_b**	**Side chain pK_a**	**Side Chain pK_b**
Arginine	R	2.17	9.04		12.48
Aspartic Acid	D	1.88	9.60	3.65	
Glutamic Acid	E	2.19	9.67	4.25	
Glycine	G	2.34	9.60		
Histidine	H	1.82	9.17		6.00
Lysine	K	2.18	8.95		10.53

Table 2.2. pK_a values of ionisable groups in proteins.

p*K*a is the negative of the logarithm of the dissociation constant of the -COOH group. pK_b is the negative of the logarithm of the dissociation constant of the -NH_3+ group.[340]

2.16 Ultra-Performance Liquid Chromatography (UPLC)

GLP-1 was reconstituted in 25 mM sodium phosphate buffer at pH 7.5 to a final peptide concentration of approximately 300 μM. The UPLC was performed using an Acquity UPLC Waters System, with a BEH C18 2.1 x 100 mm, 3.5 μm (300 Å) column attached. The method used is shown in Table 2.3.

Flow rate	**0.3 ml/min**	
Run time	40 min	
Mobile phase A	95% water, 5% acetonitrile (ACN), 0.03% trifluoroacetic (TFA)	
Mobile phase B	95% acetonitrile (ACN), 5% water, 0.03% trifluoroacetic (TFA)	
Gradient of the run		
Time (min)	Mobile phase A (%)	Mobile phase B (%)
0	75	25
30	10	90
40	0	100

Table 2.3. UPLC method.

2.17 Singular Value Decomposition (SVD)

A singular value decomposition (SVD) of far-UV CD and FT-IR data was performed using Matlab. A SVD factorises an $m \cdot n$ real matrix **M** (it is possible for a complex matrix as well but not relevant in this case) into an $m \cdot m$ real unitary matrix **U**, an $m \cdot n$ rectangular diagonal matrix with non-negative real numbers on the diagonal $\boldsymbol{\Sigma}$, and an $n \cdot n$ real unitary matrix **V**.

$$\boldsymbol{M = U \Sigma V^*} \tag{2.14}$$

The diagonal entries $\sigma_i = \Sigma_{ii}$ are the singular values of M and the columns of **U** and the columns **V** are called the left-singular vectors and the right-singular

vectors. The contribution of the different principal components of the analysed spectra was calculated by multiplying U·Σ·V*, with all values in Σ set to 0 except of $\sigma_{i,}$ as shown in Section 1.15.

2.18 Size-Exclusion Chromatography

Analytical size-exclusion chromatography (SEC) was performed by Eva Brichtová using an ÄKTA FPLC system (GE Healthcare).[341] A Superdex 200 Increase 10/300 column (GE Healthcare) was used and samples were loaded using 100, 200 or 500 µL loops. Each sample was filtered through a 0.22 µm filter prior to loading. The system was run at a flow rate of 0.75 mL min^{-1} at room temperature and UV absorbance at 280 nm was measured using a 0.5 cm flow cell. For the calibration of the column, globular protein standards (GE Healthcare) were used under the same buffer conditions as in the experiments with samples of lipidated peptides (IPP4 and IPP5).

The gel phase distribution parameter, K_{av}, of each protein standard was calculated using

$$K_{av} = \frac{V_e - V_0}{V_c - V_0} \tag{2.15}$$

Where V_e is the elution volume of a protein standard, V_0 is the void volume of the column and V_c is the total volume of the column. The logarithm of the molecular weight of each protein standard, Table 2.4, was plotted against its K_{av} and from its linear regression the molecular weight of IPP4 and IPP5 was determined. The Stoke's radius, R_{St}, was determined using a calibration plot of $(-\log(K_{av}))^{-1/2}$ against R_{St} and its linear regression, Figure 2.4. However, the commercially-available SEC calibration kits contain only globular proteins. Therefore, the values of molecular weights of oligomeric species are only estimates.

Standard	M_W [Da]	Stoke's radius [Å]	$V_{elution}$ [ml]
Blue Dextran 2000	N/A (void volume)	N/A	8.1
Aldolase	158 000	48.1	12.2
Ovalbumin	44 000	30.5	14.1
Ferritin	440 000	61	9.7
Conalbumin	75 000		13.6
Thyroglobulin	669 000	85	8.8
Ribonuclease A	13 700	16.4	17.7
Carbonic Anhydrase	29 000		16.0

Table 2.4. SEC calibration standards.

Molecular weights (MW), Stoke's radius,[342] and elution volume on a Superdex® 200 Increase 10/300 column. The calibration standards were obtained from GE Healthcare and the samples were run in 25 mM sodium phosphate buffer, pH 7.5 at a flow rate of 0.75 mL min^{-1}. Table adapted from Eva Brichtová.[341]

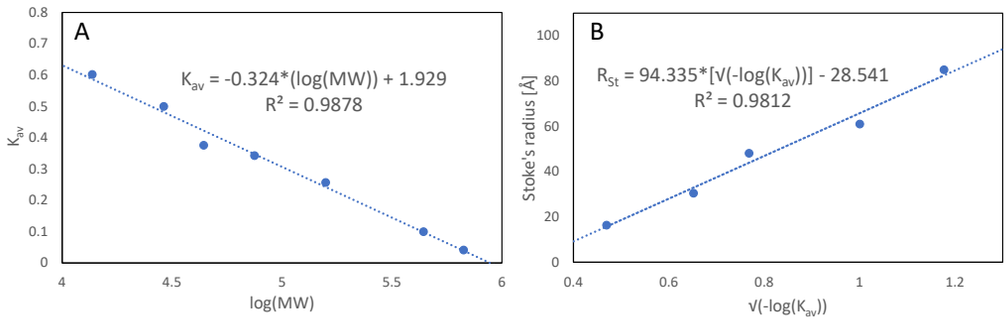


Figure 2.4. Calibration curve for SEC column

Calibration curve for the Superdex 200 Increase 10/300 column. Calibrations conducted in 25 mM sodium phosphate at pH 7.5 using proteins standards, Table 2.4. Calibration for the molecular weight (A) and the Stoke's radius (B) were performed at room temperature, with a flow rate of 0.75 mL min^{-1} and 100 µM of each protein was injected using a 200 µL injection loop. Figures adapted from Eva Brichtová.[341]

3 PHYSICAL STABILITY OF GLP-1

3.1 Introduction

Peptide therapeutics represent an increasingly important class of pharmaceutical drugs. However, despite efficacy and improvements in *in vivo* half-lives, there remain some key challenges in the development, formulation and use of these biologics. The physical stability of peptides is a common problem that, in some cases, results in major obstacles in the development pipeline of the potential drug. Poor physical stability can result in the formation of amorphous or highly structured aggregates that prevent use of the peptide or which may result in an adverse immunogenic response. The problems of physical stability are often exacerbated at the high concentrations of peptide used for subcutaneous injection. Excipients, that are commonly selected from screening experiments, are frequently used to minimise aggregation, however, they are sometimes ineffective. In order to rationally design peptide sequences/modification strategies that result in peptides with intrinsically higher degrees of physical stability, or rationally select excipients to combat aggregation, a much greater understanding of the mechanisms underpinning the aggregation of therapeutic peptides is required.

The 31-residue (7-37) therapeutic peptide GLP-1 is known to suffer from low physical stability and forms a range of aggregates under different conditions, for example, linear, rigid, amyloid-like fibrils and other species.[5] A detailed investigation of its aggregation behaviour at pH values between 7.5 and 8.2,

established that in addition to the on-pathway oligomers that can go on to form fibrillar structures, GLP-1 can form off-pathway oligomeric species at pH 7.5, which dramatically affect the kinetics of aggregation slowing down the rate of fibrillation.[5] In addition, it was shown that at all pH values investigated, freshly prepared solutions of GLP-1 contained small soluble oligomers as well as monomers. The presence of these small oligomers was particularly pronounced at pH 7.5 or 7.7 where the peptide has a higher propensity to form off-pathway oligomers and slow aggregation kinetics. In addition, the aggregation of GLP-1 has been investigated at pH 3.0 and pH 3.5, Section 1.10, and it was established that amyloid fibril formation also takes place at these acidic pH values, and that, similar to what was previously observed at pH 7.5-8.2, even small changes in pH can result in a switch in the self-assembly mechanism from predominantly on-pathway oligomers being populated to conditions where off-pathway species become populated.

In this Chapter, the aggregation of GLP-1 over a wide range of pH values and peptide concentrations is studied in further detail. In particular, the reproducibility of the aggregation assays was initially assessed. Once reproducibility had been established, time-course experiments, in which samples were taken at different times during the aggregation reaction and analysed using a range of biophysical approaches, were undertaken. Far-UV CD and FT-IR were used to provide information on the secondary structure at different times. The spectra from both techniques were subsequently analysed using a single-value decomposition method to discover more about how the structure evolved with time. AFM was used to image any species in the aggregating sample large enough to be detected using this method, providing information on higher-order structures.

The time-course experiments led to a deeper understanding of the differences between samples which showed predominantly on-pathway behaviour in comparison to samples where off-pathway oligomers were populated. The observed differences in terms of the evolution of the secondary structure, the size of the aggregates, as well as evidence for aggregates with different structures, is discussed in detail.

RESULTS

3.2 Aggregation of GLP-1 at Neutral and Basic pH

3.2.1 GLP-1 Lag Time Comparison of Data collected by KLZ and FJB

In initial experiments performed at the start of the PhD, some of the GLP-1 ThT aggregation assays that had been performed previously by Dr Karolina Zapadka (KLZ) were repeated. This was undertaken for a number of different reasons. First, to verify the results of experiments performed by KLZ, second to ensure good training in the assays, and third to check for any issues of reproducibility between different researchers.

The peptide concentration dependence measurements of the aggregation of GLP-1 were repeated at pH 7.5 and 8.2. The data are shown in Figure 3.1A & B and Figure 3.2A & B and show the expected trend where t_{lag} and $t_{1/2}$ increase with increasing peptide concentration at pH 7.5, but decrease with increasing peptide concentration at pH 8.2, as had been reported previously. The variance between measurements recorded in triplicate was, however, slightly more significant compared to the data previously obtained Dr Karolina Zapadka, Figure 1.18E & G. No additional peptide concentration dependence measurements at pH 3.0 and 3.5 were undertaken in order to save peptide from the same batch for other experiments.

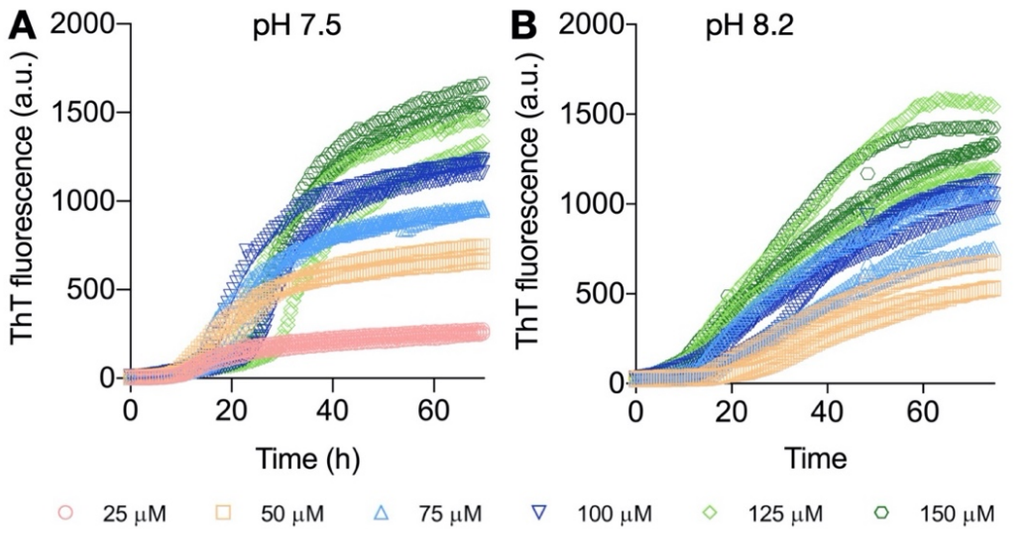


Figure 3.1. ThT assay data following the aggregation of GLP-1 at different peptide concentrations.

All samples were measured at 37 °C and at the following peptide concentrations: 25 µM (pink), 50 µM (orange), 75 µM (blue), 100 µM (dark blue), 125 µM (green), and 150 µM (dark green). A) GLP-1 aggregated in 25 mM phosphate buffer at pH 7.5 B) GLP-1 aggregated in 25 mM Tris buffer at pH 8.2.

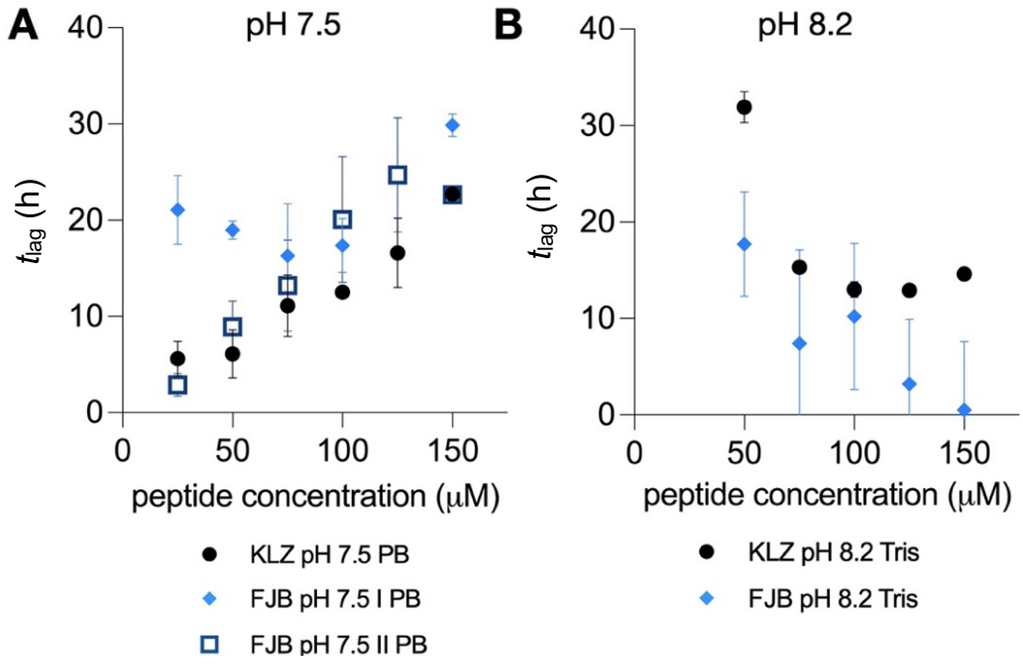


Figure 3.2. Comparison of the kinetic parameters obtained from ThT aggregation assays performed by FJB and KLZ.

Lag times of GLP-1 at different peptide concentrations. The error bars show the standard deviation. A) GLP-1 aggregated in 25 mM phosphate buffer at pH 7.5. The graph shows data previously obtained by Dr Karolina Zapadka (KLZ), and two sets of repeated triplicate measurements of GLP-1 performed by FJB, set I and set II). B) Lag times of GLP-1 aggregation in 25 mM Tris buffer at pH 8.2. The graph shows data previously obtained by Dr Karolina Zapadka (KLZ), and a triplicate measurement of GLP-1 performed by FJB.

3.2.2 Reproducibility of GLP-1 ThT Assays with Different GLP-1 Batches

One additional measurement was undertaken using a different batch (Bachem - B2) of GLP-1 which was previously used in the experiments published by Zapadka *et al*.[5] A comparison of the lag times obtained from the aggregation kinetics at pH 7.5 for the different batches is shown in Figure 3.3A. The measured values are in good agreement with the data previously obtained by Dr Karolina Zapadka[5] which shows that results of the ThT aggregation assays can be reproduced. N.B., both datasets also show the unusual increase of the lag time with increasing peptide concentration typical under these conditions. However, the differences in lag time between Batch 2 and Batch 3 are an important reminder that batch-to-batch variation must always be considered in the study of peptide aggregation, Figure 3.3B. In the rest of this Thesis, all samples of the different peptide analogues were taken from a single batch, which in the case of GLP-1 was Batch B3.

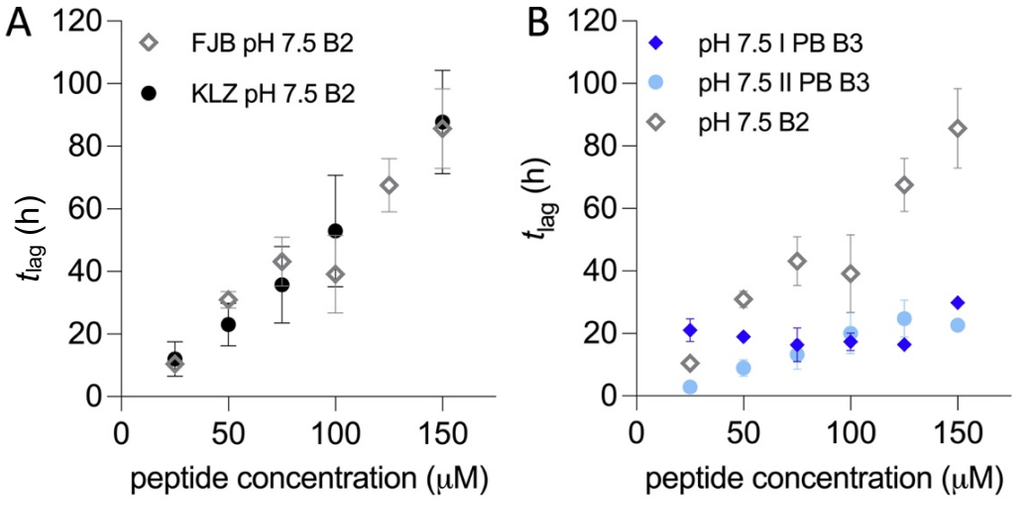


Figure 3.3. Comparison of the aggregation lag times of different batches of GLP-1 at pH 7.5.

A) Lag times for GLP-1 aggregation in 25 mM phosphate buffer at pH 7.5. The graph shows repeated measurements recorded by FJB (grey diamonds) and previously obtained data from Dr Karolina Zapadka (black solid circles). B) Lag times for GLP-1 aggregation recorded by FJB. Peptides taken from two different batches (Batch B2 and B3) and aggregated in 25 mM phosphate buffer at pH 7.5. The error bars shown in A & B correspond to the standard deviations of the kinetic parameters determined from one experiment where each peptide concentration was run in triplicate.

3.2.3 GLP-1 Aggregation Kinetics Around pH 7.5 in Tris and Phosphate Buffer

To compare the influence of different buffers, GLP-1 aggregation was measured in phosphate buffer (PB) at pH 7.5, and Tris at pH 7.2 and 7.5, all at 37 °C and at a buffer concentration of 25 mM. All measurements in 25 mM phosphate buffer, pH 7.5 show the same trend as previously reported, where the lag time increases with increasing peptide concentration. In Tris buffer at pH 7.5 (37 °C), the trend was not as pronounced and values of t_{lag} varied less, Figure 3.4, however, there was still some evidence to suggest the formation of off-pathway oligomers under these conditions. Under slightly more acidic conditions, at pH 7.2, the lag time of GLP-1 aggregation increases again with increasing peptide concentration from approximately 30 to 60 h, Figure 3.4, and it is interesting to note that it is slower than at pH 7.5

These results suggest that the identity of the buffer plays a rather small role in GLP-1 aggregation kinetics compared with the pH. The small differences that were observed most likely can be explained by the different ionic strengths of the phosphate and Tris buffers at the same pH.

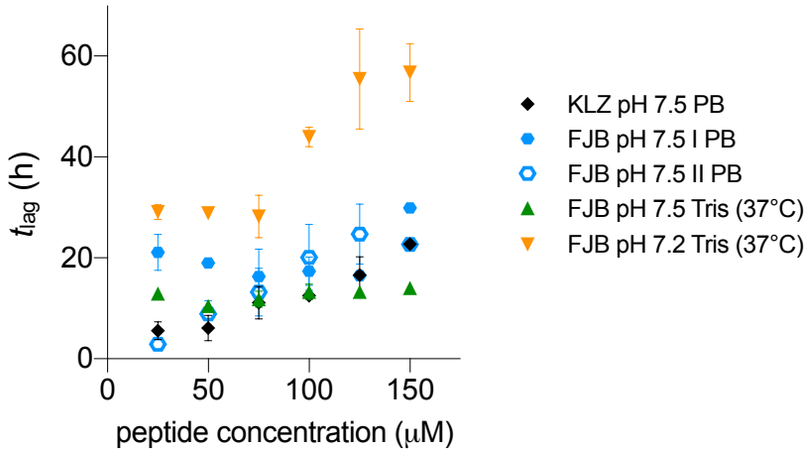


Figure 3.4. Comparison of the lag time of GLP-1 aggregation in different buffers around pH 7.5

Lag times of GLP-1 aggregation at different peptide concentrations. The graph shows data previously obtained by Dr Karolina Zapadka (black solid diamonds), as well as other measurements performed by FJB: two sets of repeated triplicate measurements in phosphate buffer, pH 7.5 (blue solid and open hexamers), triplicate measurements in Tris buffer at pH 7.5 (37 °C) (green solid triangle) and a triplicate measurement in Tris buffer at pH 7.2 (37 °C) (orange solid triangle). The error bars show the standard deviation.

3.3 Solubility and Net Charge of GLP-1

From previously conducted ThT assays on the aggregation of GLP-1, it was seen that GLP-1 aggregates not only under neutral and basic but also under acidic conditions. Additionally, the unusual pH and concentration dependence seen at pH 7.5[5] was also shown to exist at pH 3.5, Section 1.10. Understanding why such behaviour occurred under certain conditions but not all, for example, at pH 3.0 and 8.2, was one of the main motivations of this study. To provide information on how the solubility and net charge of GLP-1 vary with pH, the solubility and the pI of GLP-1 were determined experimentally, and the net charge was calculated at different pH values.

After dissolving approx. 0.2 mg of GLP-1 in approx. 50 μL of buffer at different pH values to get a final concentration of approx. 1200 μM, samples were centrifuged to get rid of undissolved peptide, and the peptide concentration determined spectrophotometrically. The results showed an apparent minimum in solubility around pH 4.5 and 5. As expected, the solubility correlates reasonably with the net charge and shows a minimum when the net charge is zero/very low. However, the pH values of the minimum solubility are slightly more acidic in comparison to both the calculated pI value 5.4 and the experimentally measured pI of 5.8 using iCIEF, Figure 3.5B & C.

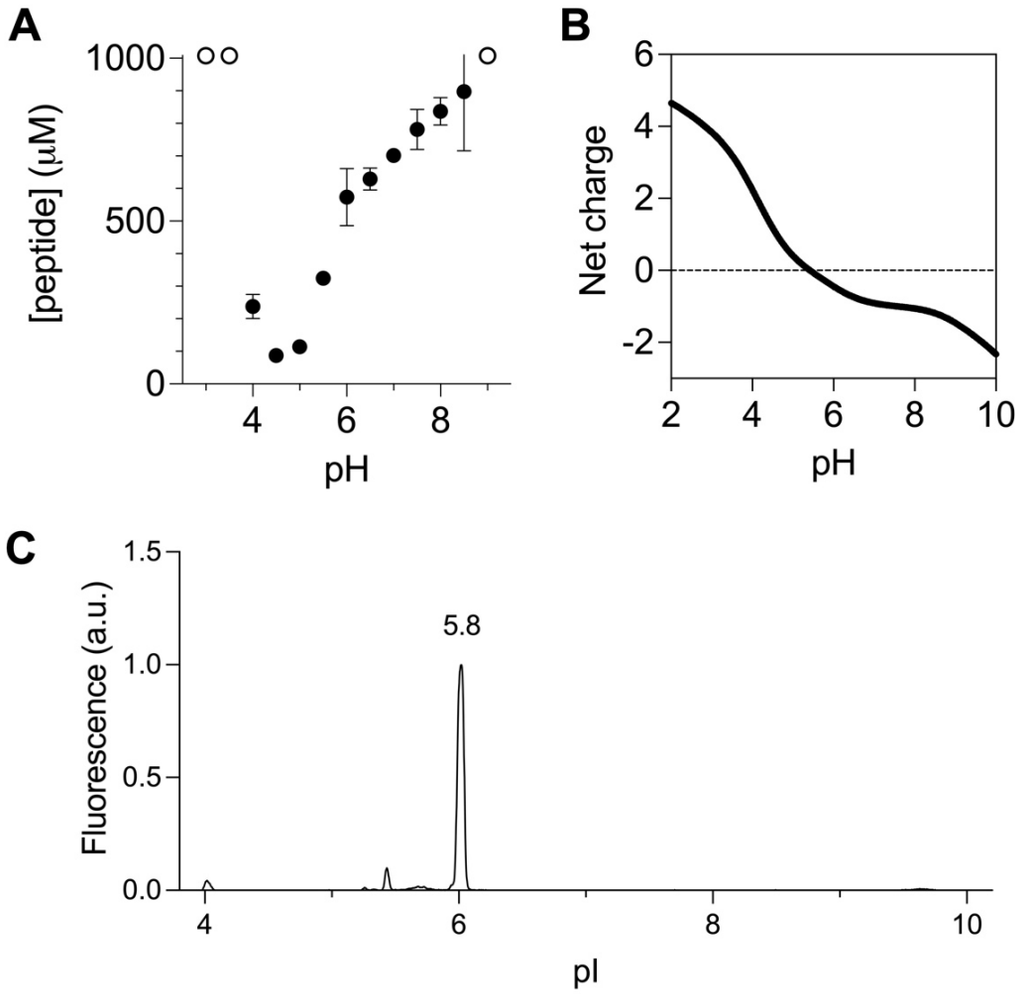


Figure 3.5. Solubility measurement, net charge calculation and pI measurement for GLP-1 using capillary isoelectric focusing.

A) Peptide concentration measured after dissolving approximately 0.2 mg GLP-1 into 50 μL buffer. The open circles indicate a concentration of at least 1000 μM. The error bars show the standard deviation of three peptide concentration measurements. B) Calculated net charge of GLP-1 peptide using the sequence and the Henderson-Hasselbalch equation, Section 2.15. C) pI measurement for GLP-1 using iCIEF.

3.4 Time-Course Experiments

In order to further investigate the mechanism by which GLP-1 aggregates to form amyloid fibrils at pH 3.0, 3.5, 7.5 and 8.2, a series of experiments were conducted aimed at characterising the size and structure of species in solution at different time points during the aggregation reaction. In particular, time points were chosen to focus on changes occurring in the lag phase of the aggregation kinetics. The size of the species formed was assessed using atomic force microscopy (AFM) and dynamic light scattering (DLS); the secondary structure assessed using far-UV CD and FT-IR. In addition, GLP-1 aggregation under acidic conditions was measured in both citrate and phosphate buffer as citrate has a strong absorbance in the FT-IR experiments in a similar region to the peptide and also leads to higher background noise in the far-UV CD spectra.

3.4.1 Imaging of Peptide Aggregates

To learn more about the size and structure of GLP-1 aggregates formed during the time course of the aggregation process AFM images were taken. A small sample was taken during a ThT assay at each time point marked with a red downwards facing triangle as shown in Figure 3.6A.

A comparison of the results obtained at different times and pH values shows that a large number of different species are populated, before, during and after aggregation depending upon conditions. These species are classified into eight different types. Curly and tangled fibrillar structures formed immediately after the monomeric peptide was dissolved are called Type s. Oligomer-like structures also observed at the start of the aggregation reaction are referred to as Type o. The most common species observed was mainly straight fibrils, however, even these are heterogeneous and have varying heights and widths. Here, the differences in height are used to distinguish six different types of fibrils. Fibrils with heights ranging from 0.5 - 1.5 nm are classified as Type a, between 1.5 - 2.5 nm as Type b, between 2.5 - 3.5 nm as Type c, between 3.5 - 5 nm as Type d, between 5 - 10 nm (Type e) and over 10 nm as Type f. The most commonly observed fibrils were Types c and e and these are likely to be

assembled from multiples of thinner fibrils. Many, even thicker fibrillar structures were observed, however, as they were normally simple multiples of the previously defined fibril types, no additional types were introduced.

N.B. The definition of the different types of fibrils does not mean that the fibrils of the same type are identical if formed under other conditions. It rather indicates that they share similar dimensions.

At pH 3.0, GLP-1 was measured in citrate and phosphate buffer. In citrate buffer no aggregates could be observed directly after the preparation and filtration of the samples, Figure 3.6B. However, after only 4 hours, thin and curvy fibrils were observed, Figure 3.6C. Additionally, some entangled aggregates can be seen embedded into the thin mesh of fibrils. In total, at least four different fibrillar species with other profiles can be seen. The thinnest fibril (Type a) has a width of approx. 15 ± 2 nm and a height of 1.2 ± 0.1 nm. The width of the second smallest fibril (Type b) shows only a slight increase to 16.7 ± 0.2 nm. However, it nearly doubles in height to 2.0 ± 0.2 nm. The width of the third (Type c) and fourth type of fibril observed (Type d) are 18 ± 1 nm and 22 ± 2 nm, respectively, and their heights are 2.8 ± 0.2 nm and 4.8 ± 0.2 nm, respectively, Figure 3.7C. After seven hours, the fibrils are less curvy, and the aggregates can be found in larger patches, Figure 3.6D. Analysis shows that after seven hours, fibrils of type c, d remain but a new fibrillar structure (Type e) appears. Interestingly, the fibrils of Type c observed after seven hours have a similar height of 2.9 ± 0.2 nm as fibrils of Type c formed after four hours but are approx. 5 nm smaller in width (13 ± 1 nm). This is also true of fibrils of Type d which, after 7 hours, also have smaller widths by approx. 7 nm. As soon as the plateau phase was reached in the ThT assay and at the end point, the mica sheet was entirely covered by a dense mesh of rigid fibrils, making it challenging to differentiate and analyse single fibrils in detail, Figure 3.6E & F.

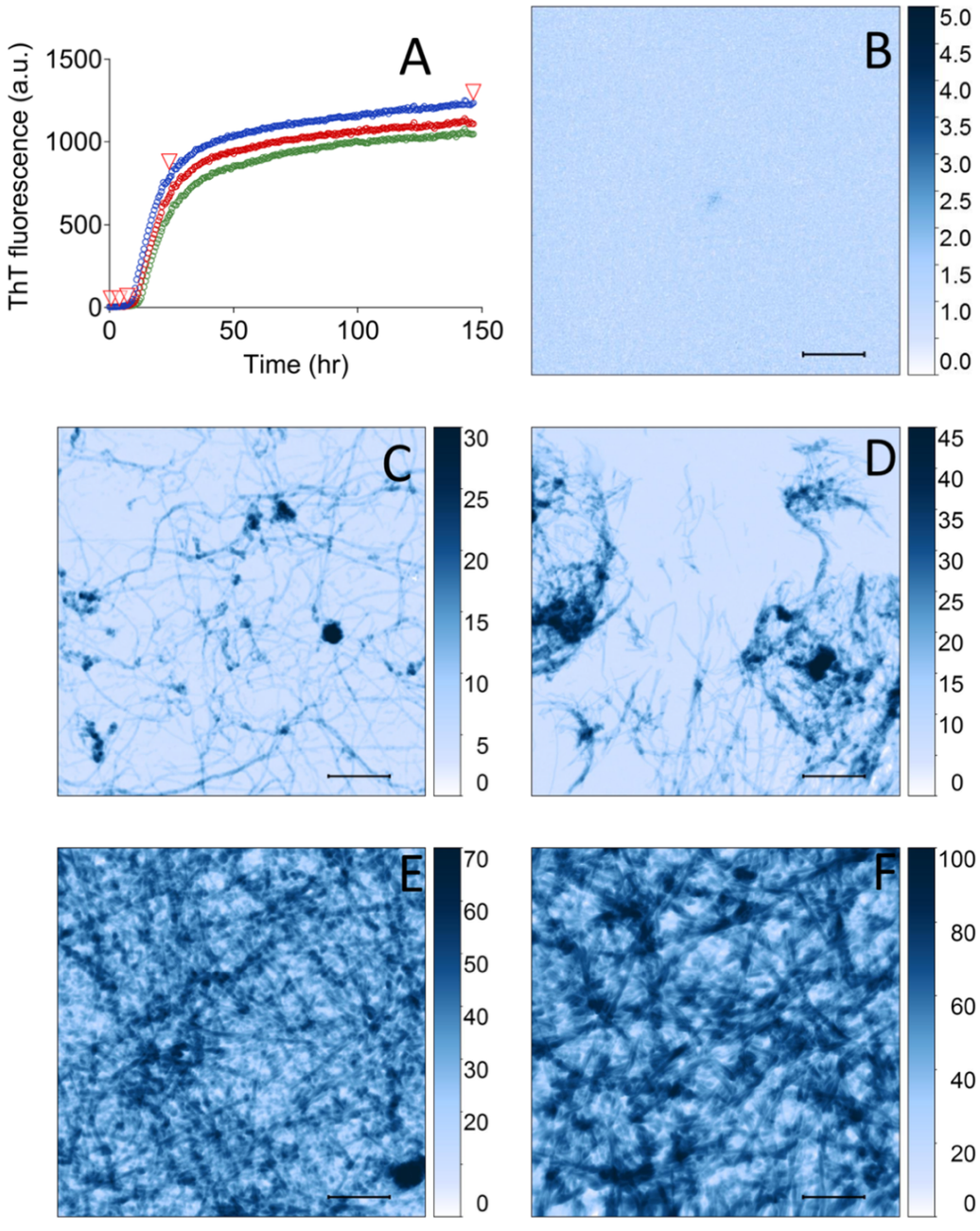


Figure 3.6. AFM images of GLP-1 species formed at different time points during peptide aggregation in citrate buffer at pH 3.0.

A) ThT assay run in triplicate of 150 µM GLP-1 in 25 mM citrate buffer at pH 3.0 and 37 °C. Triangles indicate timepoints at which AFM samples were taken. B)-F) 3x3 µm AFM pictures with inverse colour scheme (nm). Inset scale bar is 0.5 µm. The first time point was taken within one hour after preparation and filtration (B), 4 (C), 7 (D), 24 (E) and 168 (F) h.

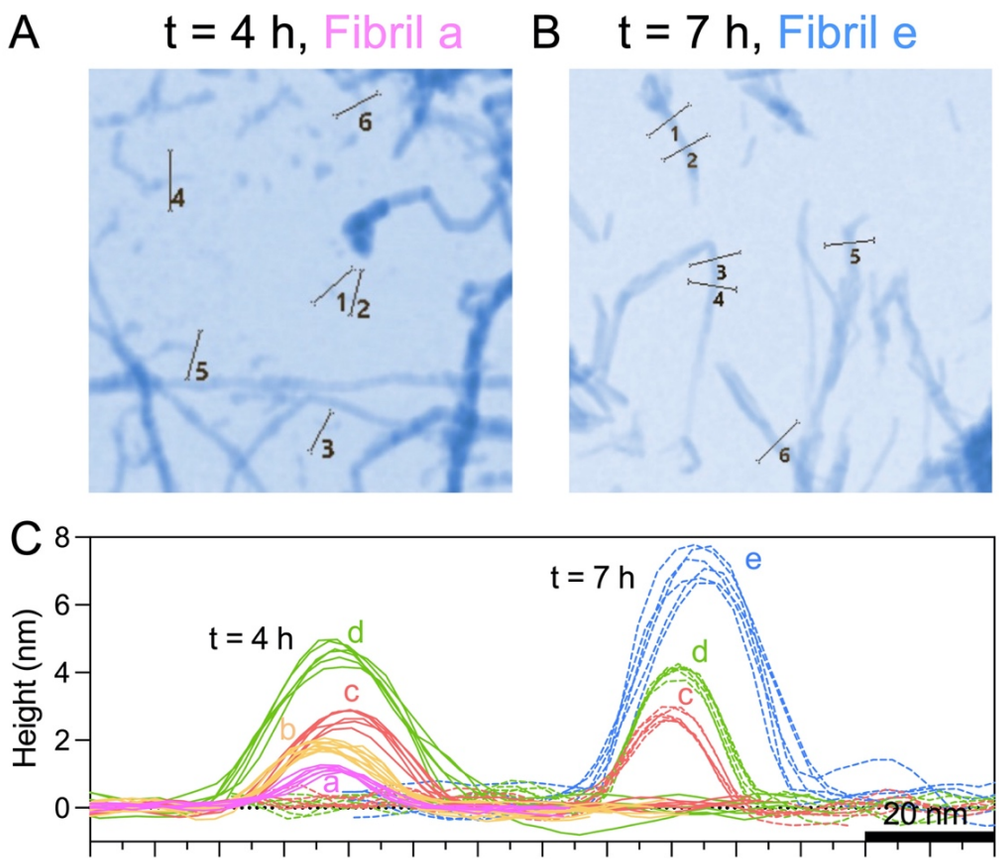


Figure 3.7. AFM images and analysis of the height and width of GLP-1 aggregates formed in citrate buffer at pH 3.0.

Measurement of height and width of fibrils shown in Figures 3.6C & D and 3.7A & B. The profiles were analysed using the AFM data analysis program Gwyddion. A) AFM image of GLP-1 aggregates (Type a) after 4 h of incubation, showing the positions on the fibrils where the profiles were measured. B) AFM image after 7 h of incubation showing thicker fibrils (Type e). Indicated are the positions where the profiles are measured. C) Complete plot of fibril profiles of different fibril types (Type a, b, c, d & e) taken after 4 h and 7 h. See Table 3.1 and Table 3.2 for further details.

GLP-1 aggregation in phosphate at pH 3.0 shows similar results to those obtained in citrate buffer: no fibrils or aggregates were observed at the earliest time point, Figure 3.8B. In the lag phase, after 2.5 and 7 hours, thin and partly curvy fibrils formed, Figure 3.8C & D. In detail, two different types of fibrils (Type c and d) were measured after 2.5 hours with heights of 2.7 ± 0.3 nm 4.1 ± 0.3 nm, respectively and similar widths between 17-18 nm, Figure 3.9A & C. After 7 hours of incubation, fibrils of Type e formed, which were slightly lower in height (5.6 ± 0.5 nm) and thicker (28 ± 2 nm) than fibrils of Type e after 2.5 hours, Figure 3.9B & C. After the steep growth phase, the surface of the mica sheet was again covered with a thick mesh of fibrils, Figure 3.8E & F.

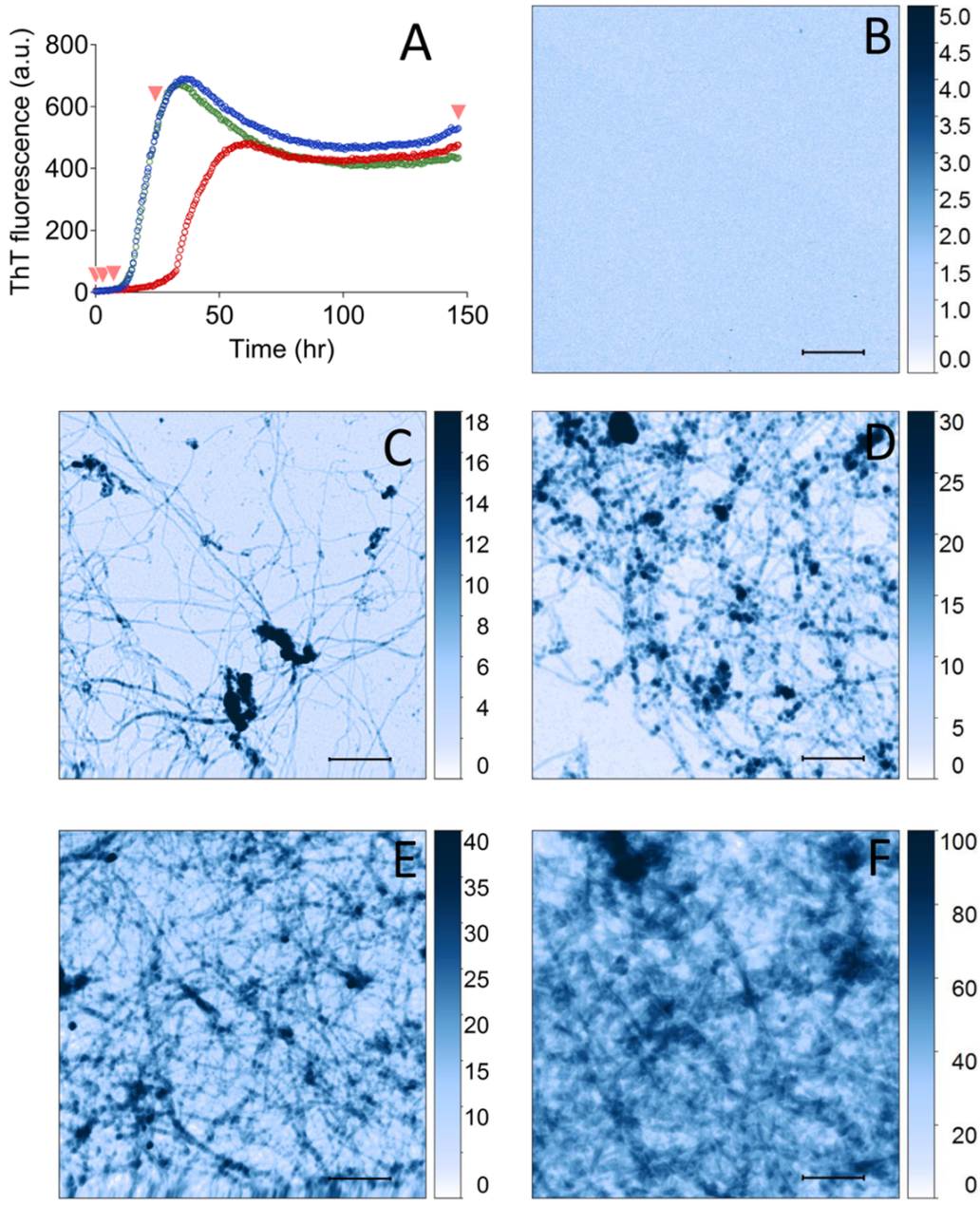


Figure 3.8. AFM images of GLP-1 species formed at different time points during peptide aggregation in phosphate buffer at pH 3.0.

A) ThT assay run in triplicate of 150 µM GLP-1 in 25 mM phosphate buffer at pH 3.0 and 37 °C. Triangles indicate timepoints at which AFM samples were taken. B)-F) 3x3 µm AFM pictures with inverse colour scheme (nm). Inset scale bar is 0.5 µm. The first time point was taken within one hour after preparation/filtration (B), as well as after 2.5 (C), 7 (D), 24 (E) and 168 (F) h.

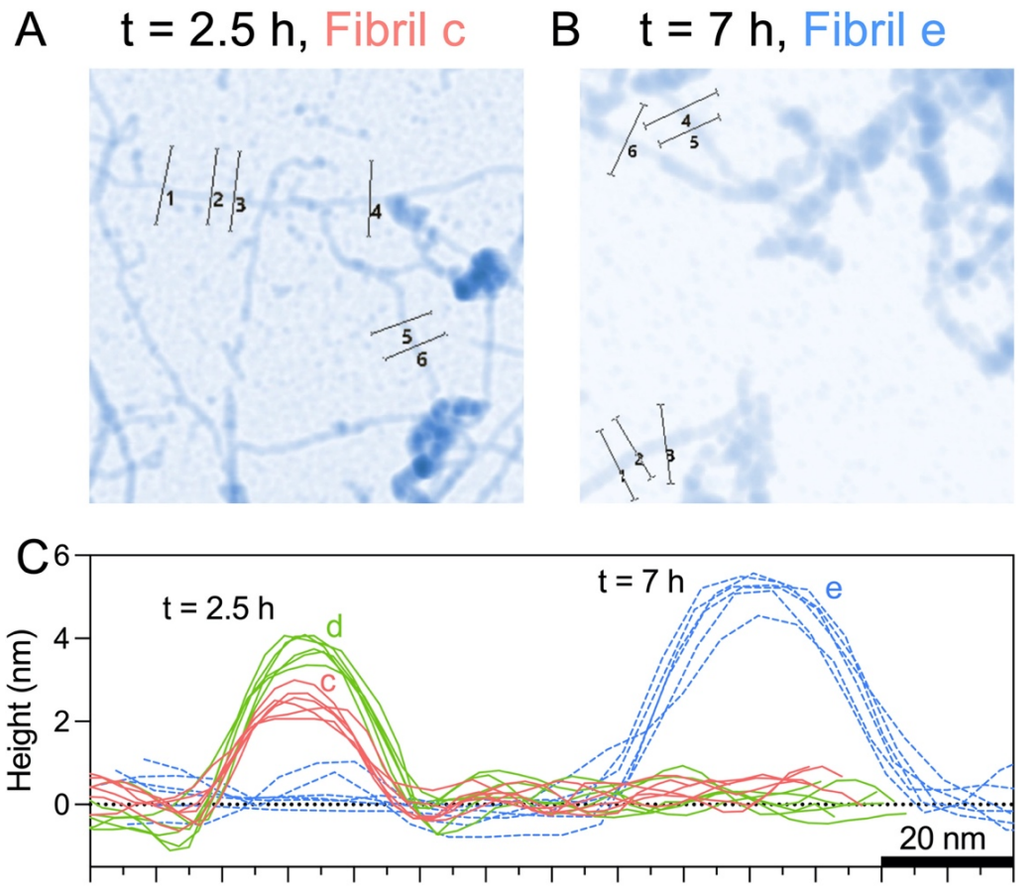


Figure 3.9. Height and width of GLP-1 aggregates formed in phosphate buffer at pH 3.0.

Measurement of height and width of fibrils shown in Figures 3.8C and 3.90A & B. The profiles were analysed using the AFM data analysis program Gwyddion. A) AFM image of sample after 2.5 h of incubation. Indicated are the places where the profiles of the thinnest fibrils (Type c) were measured. B) AFM image showing a sample after 7 h of incubation. Indicated are the six positions on fibrils of Type e where a profile was measured. C) Complete plot of fibril profiles of different fibril types (Type c, d & e) taken after 2.5 h and 7 h. See Table 3.1 and Table 3.2 for further details.

The AFM imaging of GLP-1 at different time points during aggregation in citrate buffer at pH 3.5 showed curly and tangled fibril-like aggregates immediately after the sample preparation, Figure 3.10B. The height and width of these short fibril-like aggregates was measured as 2.2 ± 0.1 nm and 15 ± 1 nm, respectively (Type s), Figure 3.11A. Some of these aggregates even formed closed ring-like structures. After 6 hours, the time point corresponding to the end of the growth phase, the sample formed fibrils with three different profiles, Figure 3.10C, Type a, b and e. The thinnest fibril (Type a) has a height of 1.5 ± 0.2 nm whilst Type b heights were 3.0 ± 0.2 nm, they both had similar widths of approx.14 nm. Type a and b fibrils are probably the predecessor for some

longer, straight, and partly twisted fibrils (Type e), Figures 3.10C and 3.11E. It is interesting to note that the width of the latter decreases from approx. 23 nm to less than 18 nm during the lag phase. After 23 hours, Type a fibrils disappear and only mostly straight, longer and more rigid fibrils were detected, Figures 3.10D-F and 3.11E. In Figure 3.12D, the tip of a thick fibril of Type f is analysed which appears to be assembled out of thinner fibrils of Type c and e.

The findings for GLP-1 at pH 3.5 in phosphate buffer show similar aggregates as in citrate buffer. Again, short and curly fibril-like structures (Type s) are formed immediately after the preparation of the samples, Figure 3.12B. The formation of ring-like structure also seems independent of the buffer. It is worth highlighting that similar as for the citrate buffer at pH 3.5 also in phosphate buffer, the thinnest fibrils (Type b) measured at the second time point are lower and thinner in comparison to the earlier formed curly and tangled fibril-like aggregates. However, during the lag phase and at the beginning of the growth phase, the images show more amorphous structures, Figure 3.12C & D. At the beginning of the plateau and at the end of the aggregation assay, the fibrils are rigid and similar to those observed in citrate buffer, Figure 3.12E & F

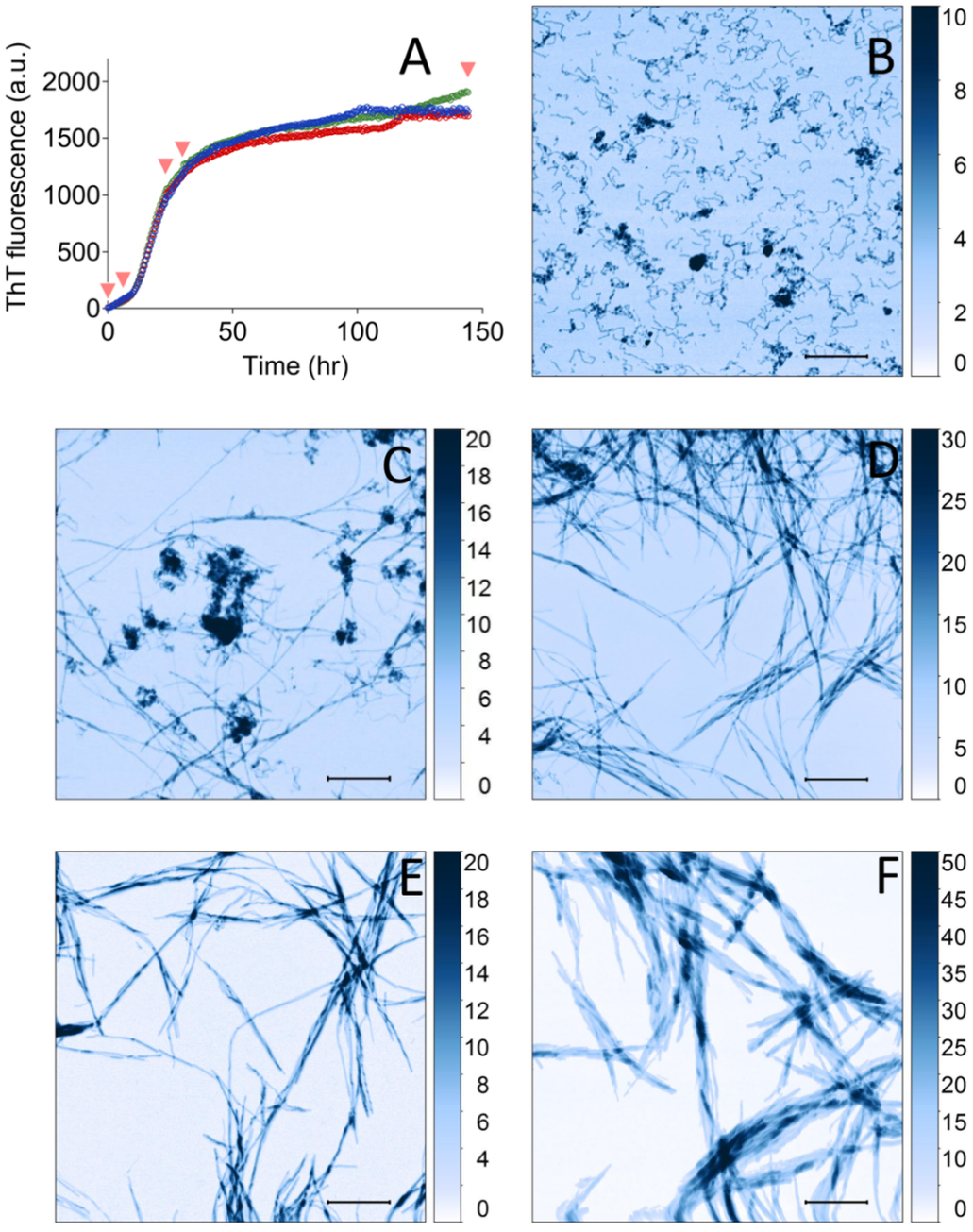


Figure 3.10. AFM images of GLP-1 species formed at different time points during peptide aggregation in citrate buffer at pH 3.5.

A) ThT assay run in triplicate using 150 µM GLP-1 in 25 mM citrate buffer at pH 3.5. Triangles indicate timepoints at which AFM samples were taken. B)-F) 3x3 µm AFM pictures with inverse colour scheme (nm). Inset scale bar is 0.5 µm. The first time point was taken within one hour after preparation/filtration (B), as well as after 6 (C), 23 (D), 30 (E) and 144 (F) h. Note, sample shown in F was diluted to ~50 µM before imaging.

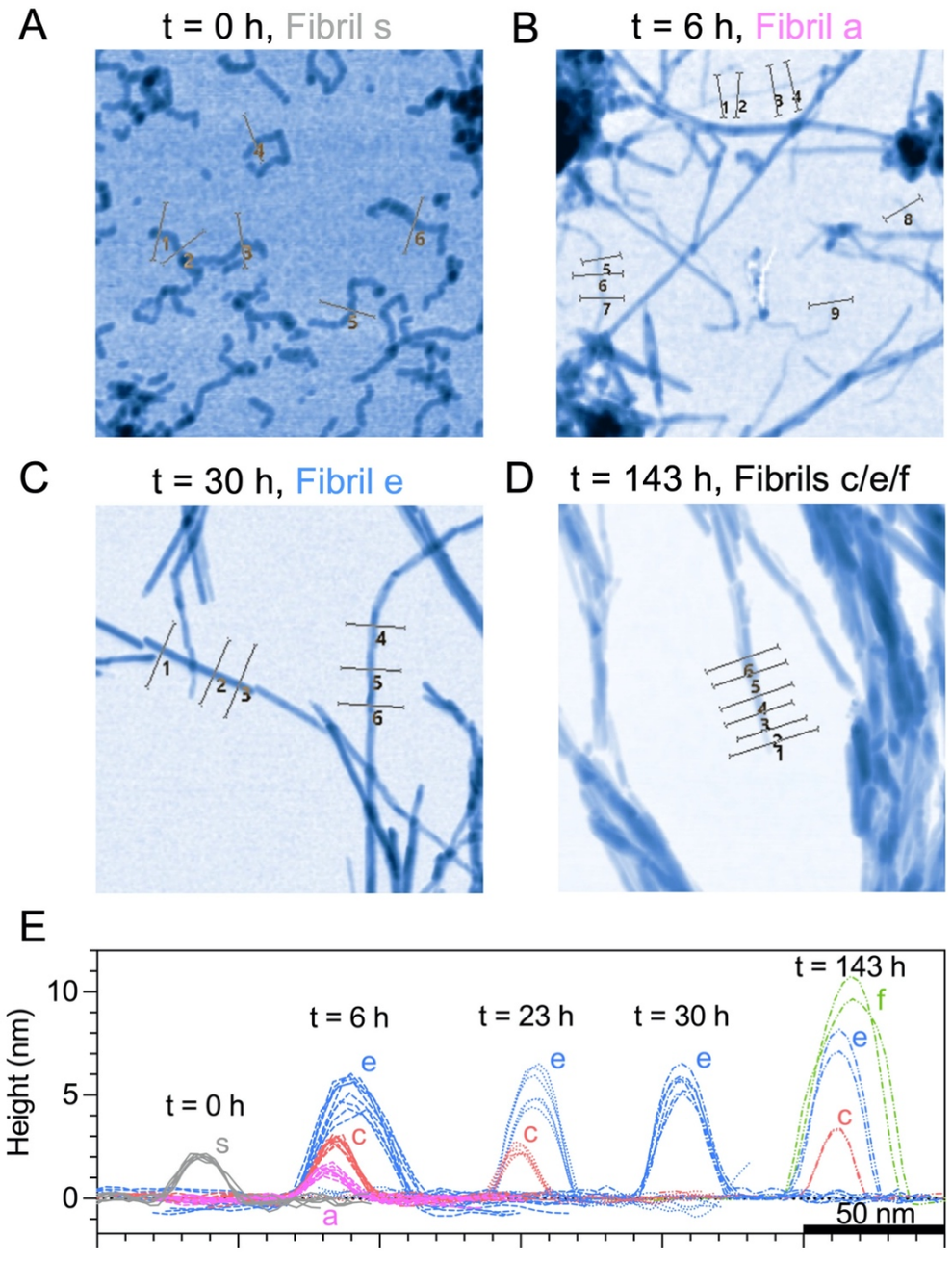


Figure 3.11. AFM images and analysis of the height and width of GLP-1 aggregates formed in citrate buffer at pH 3.5.

AFM images of GLP-1 aggregates formed in citrate at pH 3.5 illustrating the sections of the fibrils analysed. The profiles were analysed using the AFM data analysis program Gwyddion. A) Short curly fibrils (Type s) imaged before the incubation at t = 0 h. Indicated are the six positions on the fibrils where a height/width profile was measured. B) AFM image recorded after 6 h indicating the nine places where profile measurements were taken on Type a fibrils. C) AFM image after 30 h of incubation indicating the six positions on the Type e fibrils where profiles were measured. D) AFM image after 143 h showing Type d, e and f fibrils and the six positions on the fibrils where profiles were measured. E) Complete plot of fibril profiles of different fibril types (Type s, a, c, e & f) taken between 0 and 143 h. See Table 3.1 and Table 3.2 for further details.

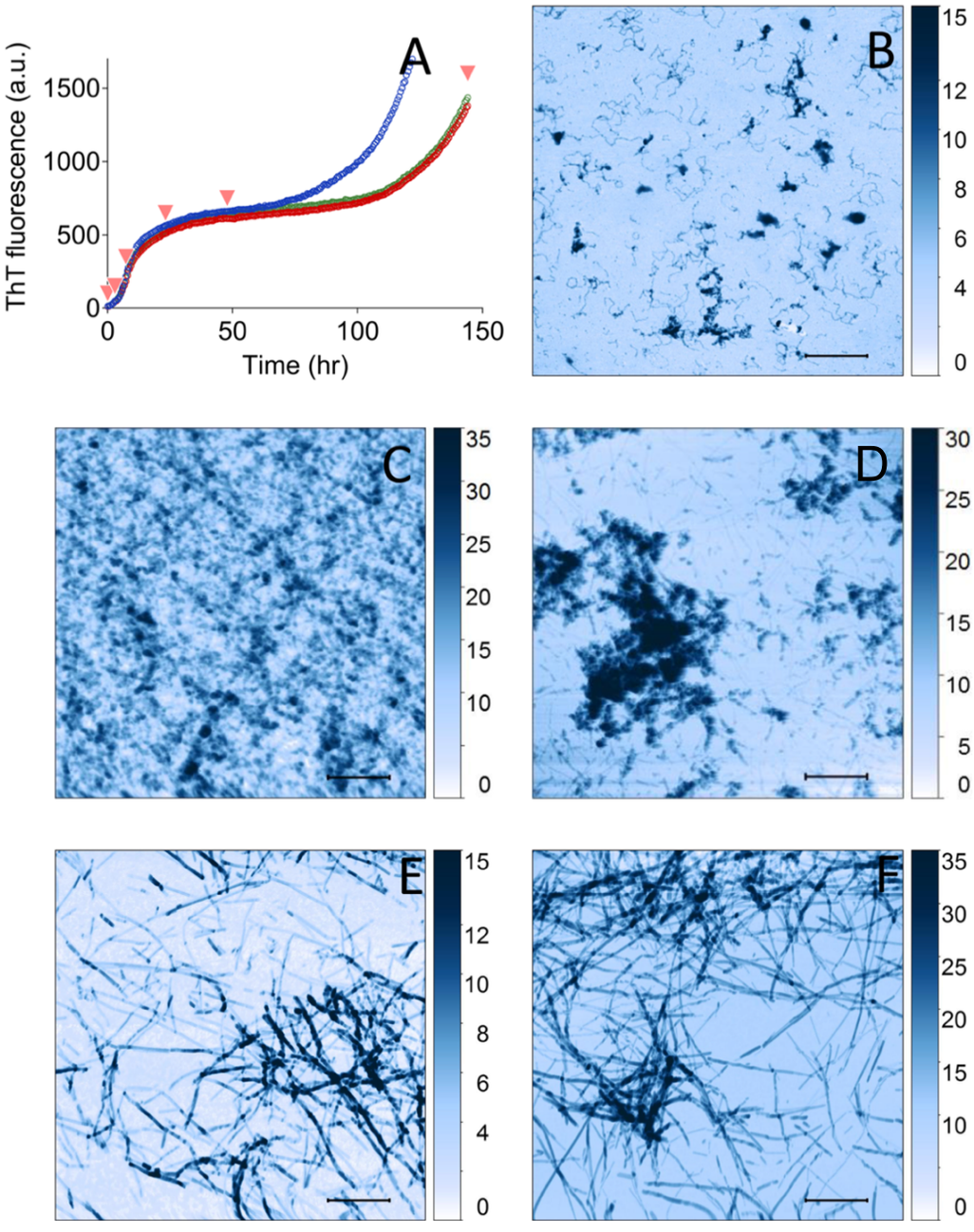


Figure 3.12. AFM images of GLP-1 species formed at different time points during peptide aggregation in phosphate buffer at pH 3.5.

A) ThT assay run in triplicate using 150 µM GLP-1 in 25 mM phosphate buffer at pH 3.5 and at 37 °C. Triangles indicate timepoints at which AFM samples were taken. B)-F) 3x3 µm AFM pictures with inverse colour scheme (nm). Inset scale bar is 0.5 µm. The first time point was taken within one hour after preparation/filtration (B), as well as after 3 (C), 8 (D), 48 (E) and 143 (F). Samples in (F) were diluted to ~50 µM before imaging.

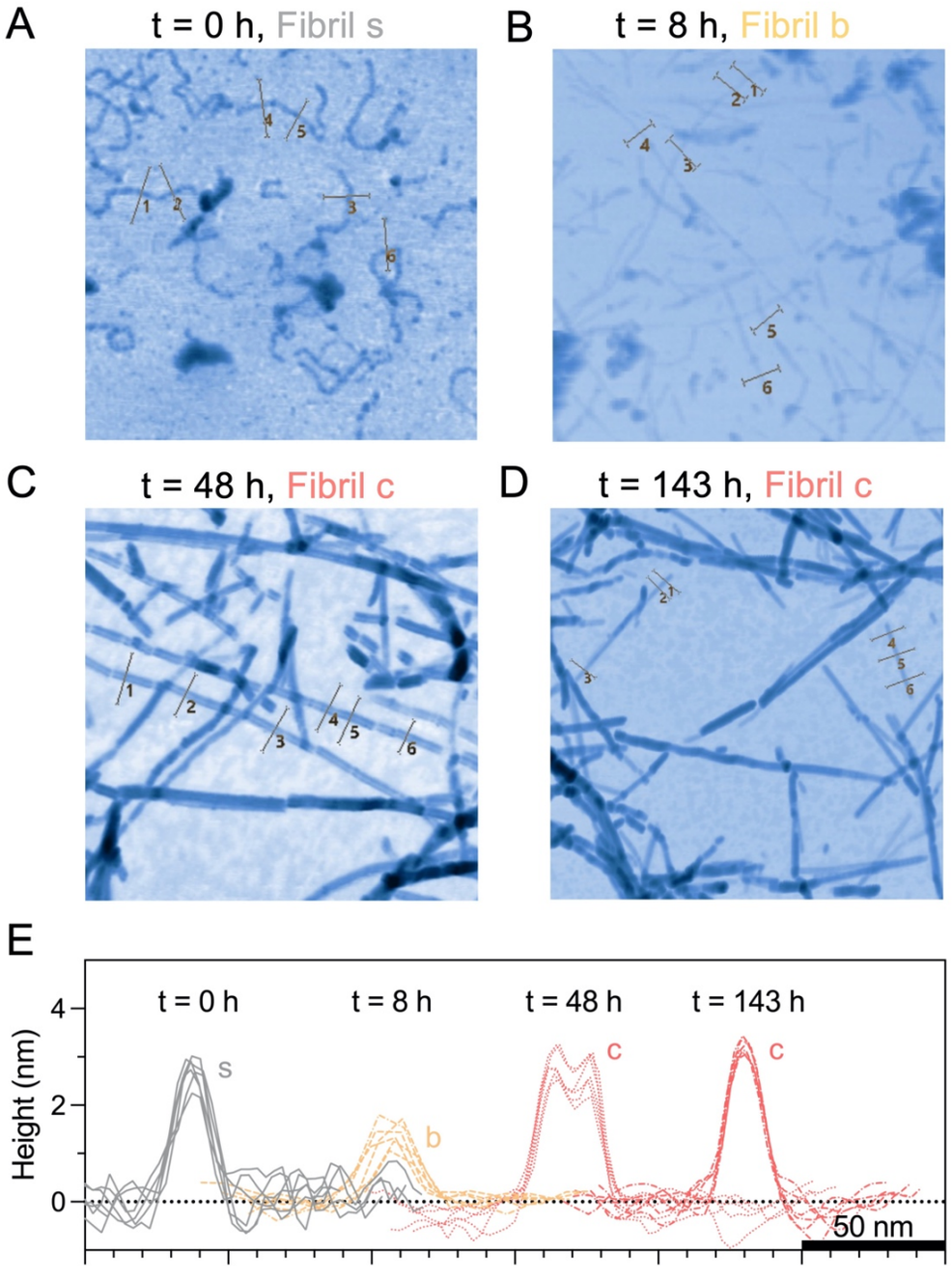


Figure 3.13. AFM images and analysis of the height and width of GLP-1 aggregates formed in phosphate buffer at pH 3.5.

Measurement of height and width of fibrils shown in Figures 3.12B, D, E & F and 3.13A-D. The profiles were analysed using the AFM data analysis program Gwyddion. A) AFM image of short curly fibrils (Type s) imaged before incubation. The six positions used to calculate the height/width of the fibrils are shown. B) AFM image of aggregates after 6 h of incubation (Type b). Shown are the six positions used to measure the height/width of the fibrils. C) AFM image of aggregates (Type c) after 48 h of incubation. The six positions used to measure the height/width of the fibrils are shown. An AFM artefact most likely causes the double peak. D) AFM image after 143 h incubation (Type c). The six positions used to measure the height/width of the fibrils are shown. E) Complete plot of fibril profiles of different fibril types (Type s, b, & c) taken between 0 and 143 h. See Table 3.1 and Table 3.2 for further details.

It is interesting to note that the ThT assay used to assess the aggregation of the samples of GLP-1 in phosphate buffer at pH 3.5 for the AFM experiments shows some differences in lag time compared to previous measurements in phosphate buffer at pH 3.5. The differences are most likely caused by the disturbance of the aggregation process resulting from the opening and closing of the plate reader during the sampling, as well as inserting a pipette to remove sample.

The aggregation of two samples of GLP-1 was also measured at pH 6.5, a pH close to the pI of GLP-1 (Figure 3.5C) and, in this case, no increase in ThT fluorescence even after the 140 h of incubation was observed, Figure 3.14A. AFM images were taken at the beginning and at the end of ThT assay to confirm that fibrils had not formed, Figure 3.14B & C.

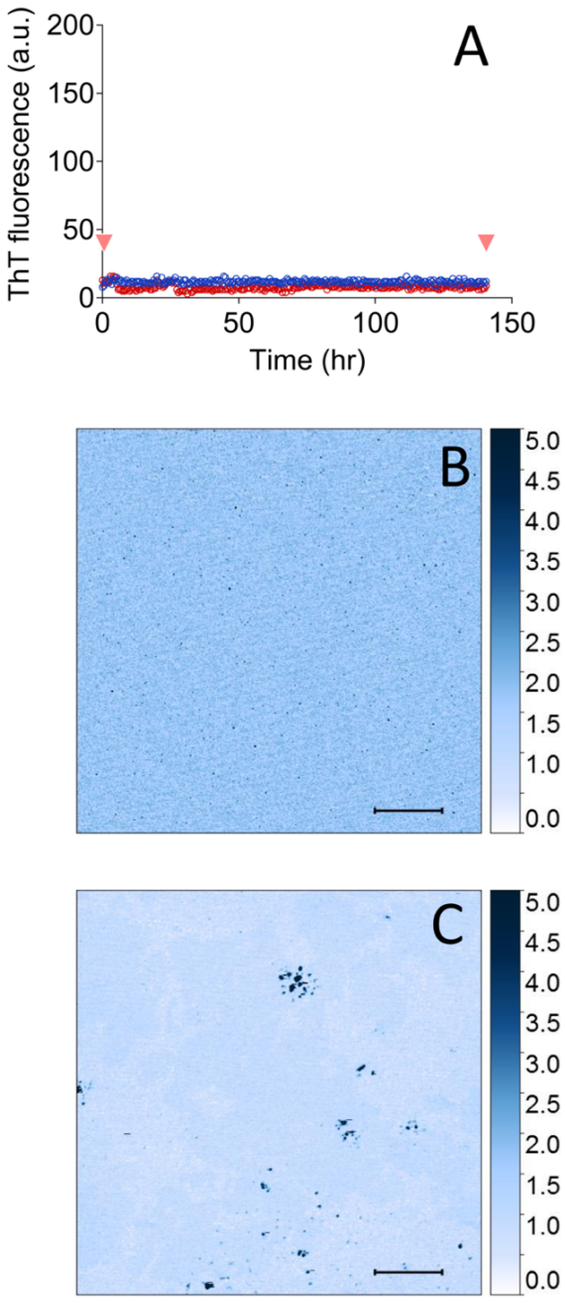


Figure 3.14. AFM images of GLP-1 species formed at different time points during peptide aggregation in phosphate buffer at pH 6.5.

A) ThT assay run in duplicate of 150 µM GLP-1 in 25 mM phosphate buffer at pH 6.5 and 37 °C. Triangles indicate timepoints at which AFM samples were taken. B and C) 3x3 µm AFM pictures with inverse colour scheme (nm). Inset scale bar is 0.5 µm. The AFM was taken within one hour after sample/preparation filtration (B) and after 144 (C) hours.

At pH 7.5, immediately after sample preparation, samples of GLP-1 contained small, thin and curly fibril-like structures (Type s), Figure 3.15B. These had heights of 1.2 ± 0.1 nm and widths of 12 ± 1 nm, and were smaller than their curly counterparts under acidic conditions, Figure 3.16A. At the end of the lag phase, the thin curvy structures convert to short and straight fibrils, Figure 3.15C. They have profiles that are typical of fibrils of Type b, Figure 3.16B. The short and straight fibrils of Type b most likely grow and form the mature fibrils of Type c to f, which are present in the samples taken from the midpoint of the growth phase, at the start of the plateau phase and at the end of the aggregation assay, Figure 3.15D-F.

At the start of the aggregation reaction, at pH 8.2, small oligomeric species (Type o) were observed with diameters of 33 ± 3 nm and heights of 1.9 ± 0.2 nm, Figures 3.17B and 3.18A. After 5 hours, these species vanish and short fibrils (Type b) appear, Figures 3.17C and 3.18B. The distribution and pattern of the aggregates suggests that this species is the predecessor of the short fibrils and, therefore, an on-pathway species, Figure 3.18A & B. After an additional four hours, these short fibrils grew only moderately in terms of height whilst their widths decreased and a few longer fibrils appeared, Figure 3.17D. The fibrils observed in the growth phase and in the plateau appear region appeared rigid, Figure 3.17E & F, and were similar to the fibrils observed at late time points for GLP-1 at pH 3.0, 3.5 and 7.5

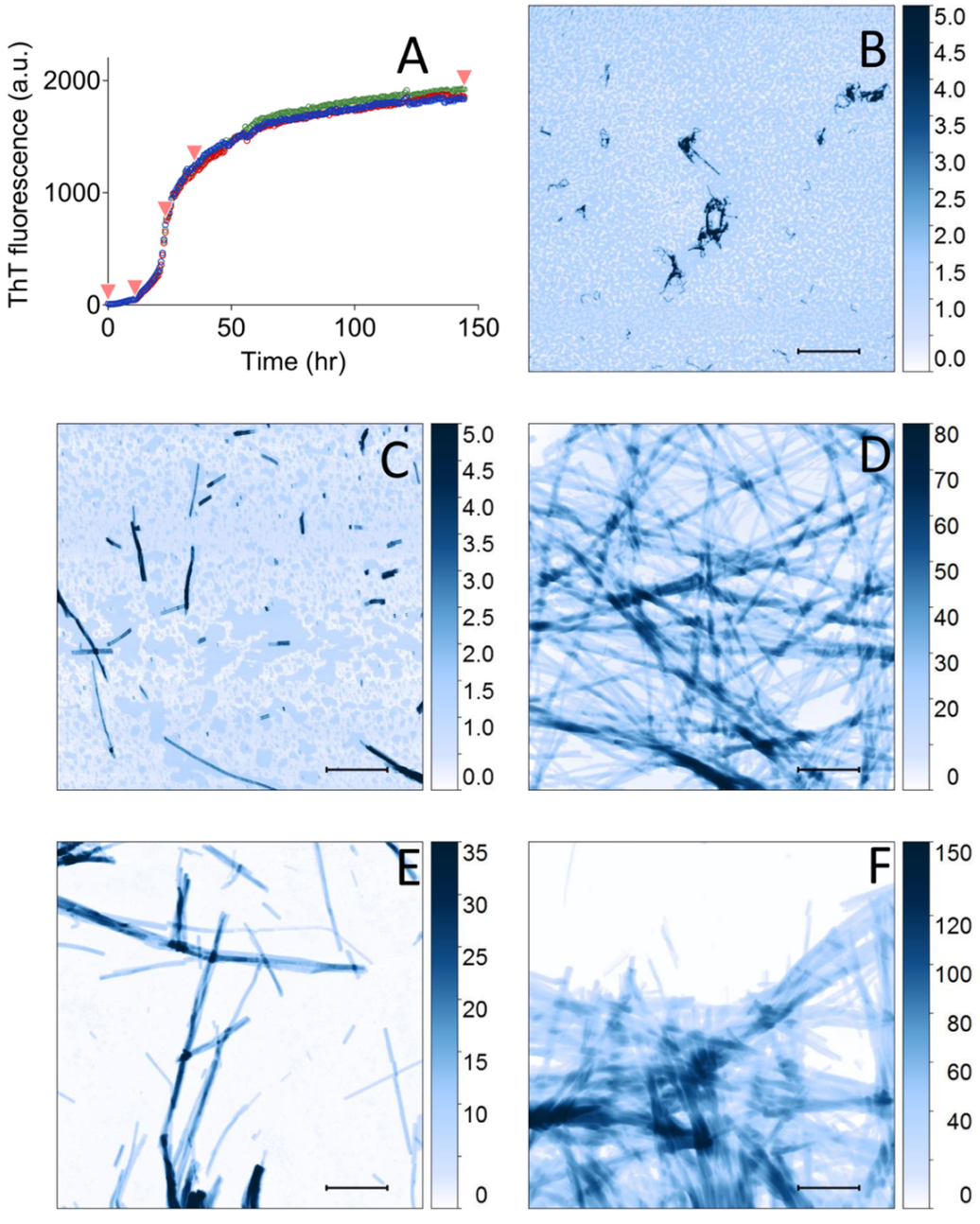


Figure 3.15. AFM images of GLP-1 species formed at different time points during peptide aggregation in phosphate buffer at pH 7.5.

A) ThT assay run in triplicate using150 µM GLP-1 in 25 mM phosphate buffer at pH 7.5 and 37 °C. Triangles indicate timepoints at which AFM samples were taken. B)-F) 3x3 µm AFM pictures with inverse colour scheme (nm). Inset scale bar is 0.5 µm. The first time point was taken within one hour after preparation/filtration (B), as well as after 11 (C), 23 (D), 35 (E) and 145 (F) h.

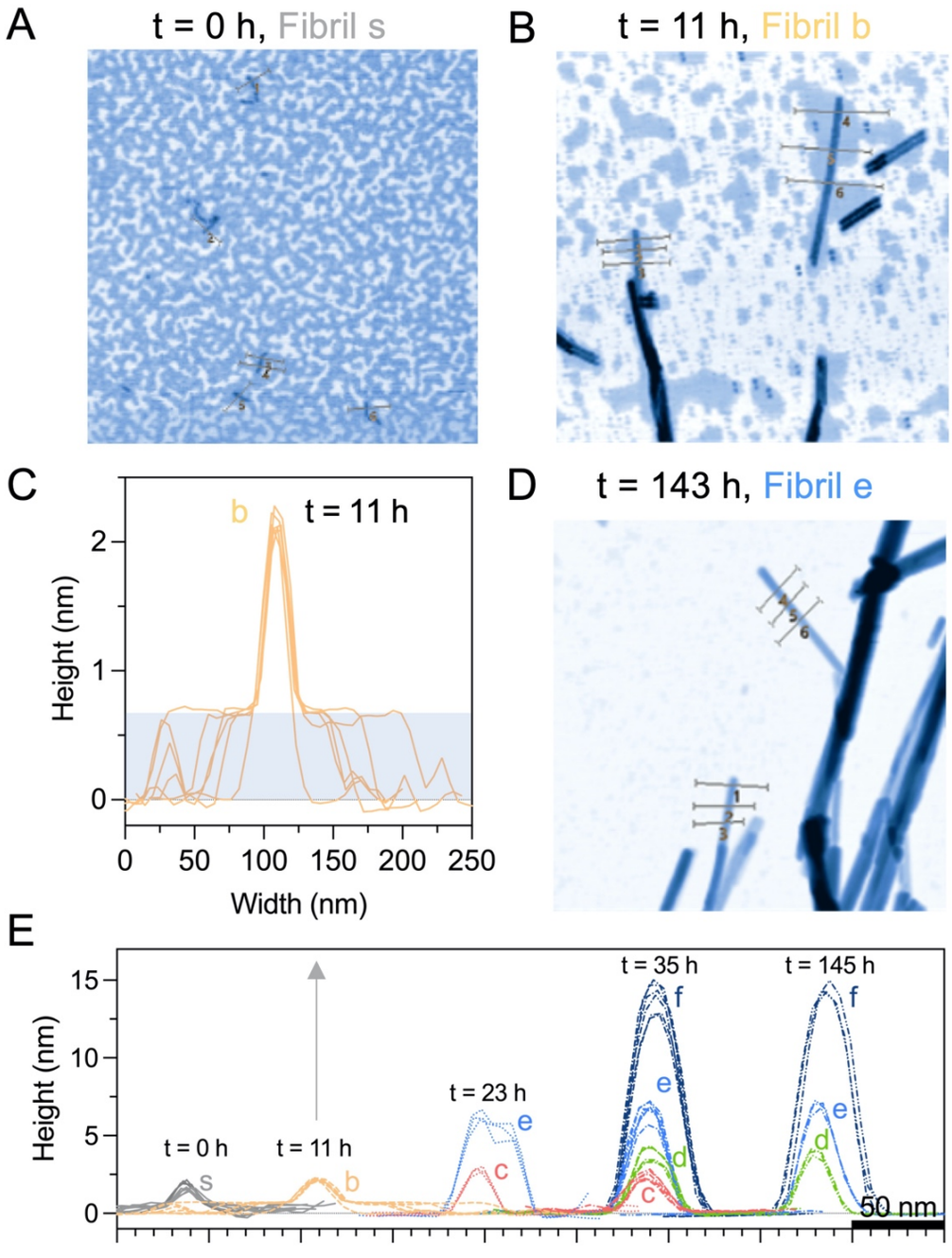


Figure 3.16. AFM images and analysis of the height and width of GLP-1 aggregates incubated in phosphate buffer at pH 7.5.

Measurement of height and width of fibrils shown in Figures 3.15B-F and 3.16A, B & D. The profiles were analysed using the AFM data analysis program Gwyddion. A) AFM image of aggregates of GLP-1 before incubation (Type s). The six positions where profiles were measured are shown. B) AFM image after 11 h of incubation (Fibril b). The six positions where profiles were measured are shown. (C) Plot of the height/width profiles of fibrils measured using data shown in panel (B). The height/width of the fibrils was measured above the blue shaded plateau. D) AFM image of aggregates formed after 143 h incubation. The six positions where profiles were measured are shown. E) Complete plot of fibril profiles of different fibril types (Type s, b, c, d, e & f) taken between 0 and 145 h. See Table 3.1 and Table 3.2 for further details.

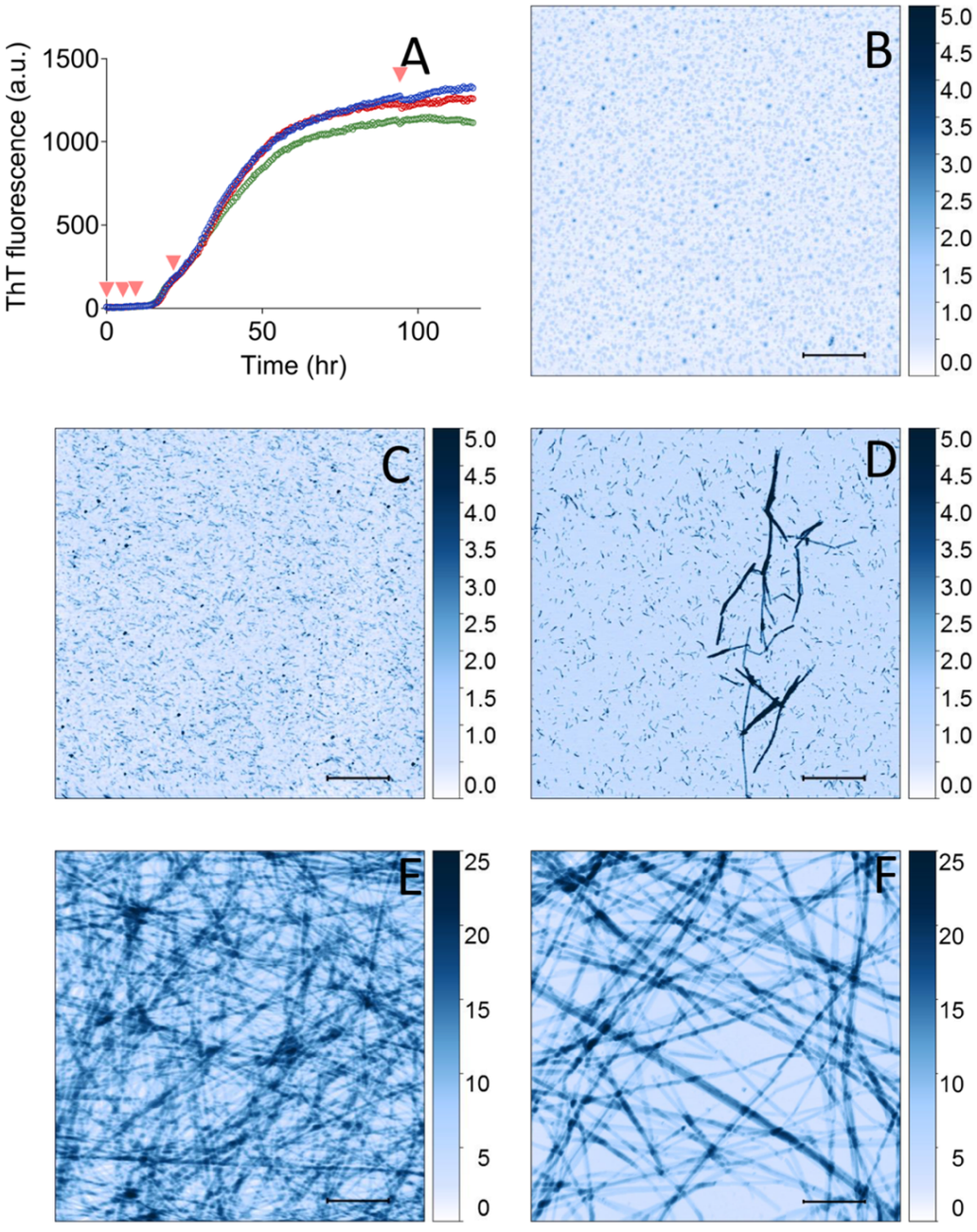


Figure 3.17. AFM images of GLP-1 species formed at different time points during peptide aggregation in Tris buffer at pH 8.2.

A) ThT assay run in triplicate of 150 µM GLP-1 in 25 mM Tris buffer at pH 8.2. Triangles indicate timepoints at which AFM samples were taken. B)-F) 3x3 µm AFM pictures with inverse colour scheme (nm). Inset scale bar is 0.5 µm. The first time point was taken within one hour after preparation/filtration (B), as well as after 5 (C), 9 (D), 21 (E) and 94 (F) h.

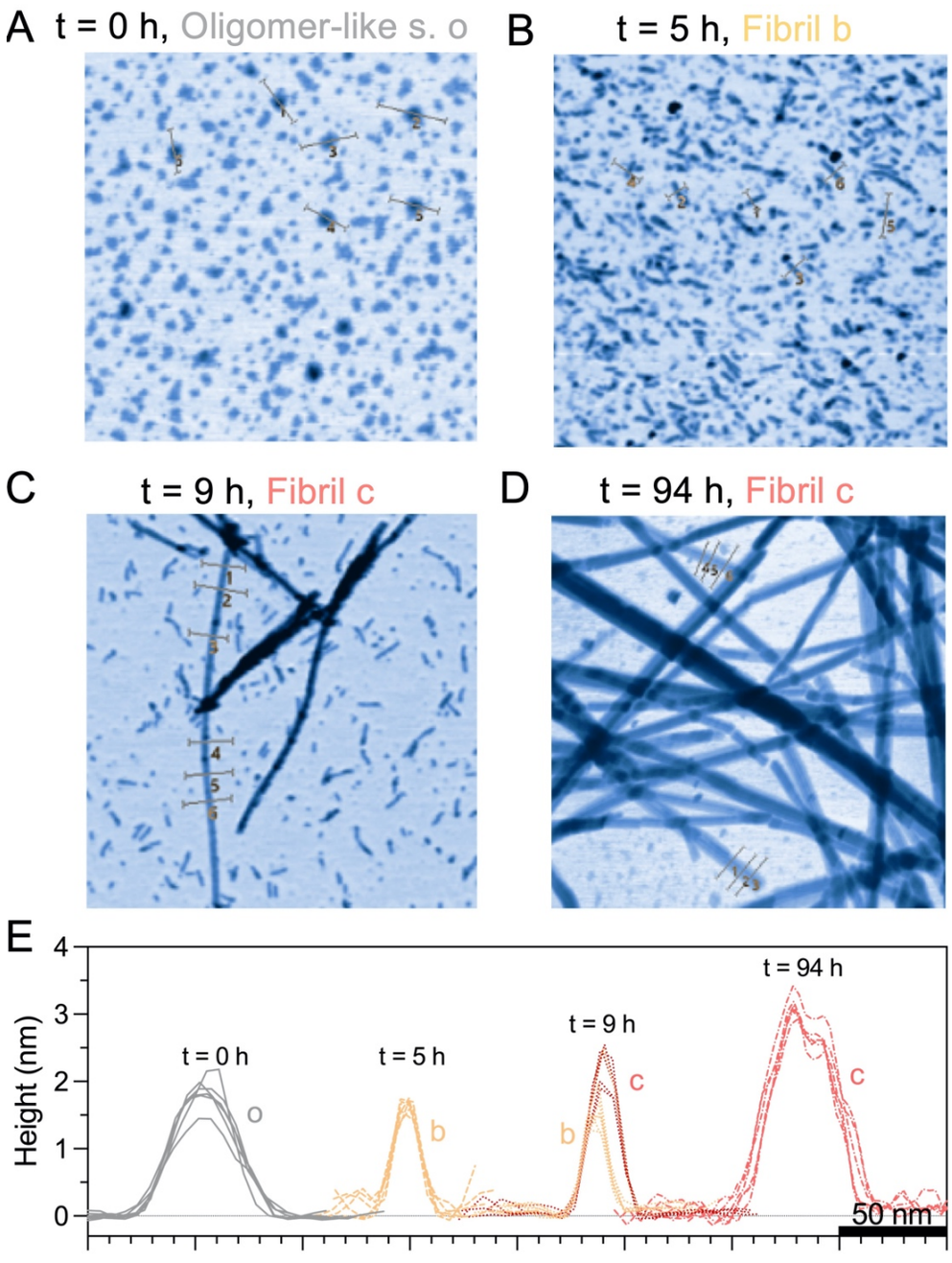


Figure 3.18. AFM images and analysis of height and width of GLP-1 aggregates incubated in Tris buffer at pH 8.2.

Measurement of height and width or diameter of small oligomers and fibrils shown in Figures 3.17B, C, D & F and 3.18A-D. The profiles were analysed using the AFM data analysis program Gwyddion. A) AFM image of GLP-1 aggregates (Type o) before the incubation. The six positions where profiles were measured are shown. B) AFM image after five h of incubation showing fibrils of Type b. The six positions where profiles were measured are shown. C) AFM image of intermediate fibrils imaged after nine h of incubation (Type c). The six positions where profiles were measured are shown. D) AFM image of thick fibrils (Type c) observed after 94 h incubation. The six positions where profiles were measured are shown. E) Complete plot of fibril profiles of different fibril types (Type o, b, & c) taken between 0 and 94 h. See Table 3.1 and Table 3.2 for further details.

Time point (h)	Fibril type	Height (nm)	Height SD (nm)	Width (nm)	Width SD (nm)
Citrate buffer pH 3.0					
4	a	1.2	0.1	14.9	1.6
4	b	1.9	0.2	16.7	1.2
4	c	2.8	0.2	17.9	1.2
4	d	4.8	0.2	21.8	1.6
7	c	2.9	0.2	12.9	0.7
7	d	4.3	0.1	14.0	0.7
7	e	7.6	0.5	19.4	1.8
Phosphate buffer pH 3.0					
2.5	c	2.7	0.3	17.3	2.1
2.5	d	4.1	0.3	17.9	2.0
7	e	5.6	0.5	27.5	2.1
Citrate buffer pH 3.5					
0	s	2.2	0.1	15.4	1.3
6	a	1.5	0.2	13.7	1.3
6	c	3.0	0.2	14.6	0.5
6	e	5.6	0.7	23.4	1.5
23	c	2.5	0.3	12.9	1.6
23	e	5.9	1.0	17.0	0.8
30	e	6.0	0.6	17.6	0.6
143	d	3.6	0.0	12.4	0.2
143	e	8.1	0.8	17.7	0.2
143	f	10.8	0.8	24.3	2.9
Phosphate buffer pH 3.5					
0	s	2.8	0.4	14.3	3.3
8	b	1.6	0.3	16.6	1.2
48	c	3.0	0.4	21.1	0.7
143	c	3.4	0.1	13.6	0.5
Phosphate buffer pH 7.5					
0	s	1.2	0.1	12.1	1.2
0	s	1.7	0.1	11.1	0.9
11	b	1.6	0.1	16.2	1.1
11	b	1.6	0.4	26.2	1.0
23	c	3.0	0.2	16.2	0.2
23	e	6.6	0.4	29.9	1.1
35	b	2.5	0.2	20.9	2.6
35	d	4.1	0.5	22.5	0.5
35	e	7.2	0.6	21.2	0.3
35	f	15.0	1.1	27.0	1.1
143	d	4.2	0.2	18.2	0.7
143	e	7.3	0.3	21.1	0.6
143	f	15.4	0.4	27.7	0.3
Tris buffer pH 8.2					
0	o	1.9	0.2	33.0	2.5
5	b	1.8	0.1	14.3	1.2
9	b	1.7	0.2	12.0	0.4
9	c	2.4	0.3	15.4	0.2
94	c	3.1	0.2	35.1	4.0

Table 3.1. Heights and widths of different fibril profiles.

The values for the height and width were determined from the fibril profiles using a Gaussian fit. The width of the fibril is equivalent to the FWHM of the Gaussian.

	pH 3.0	**pH 3.5**	**pH 7.5**	**pH 8.5**
Buffer Data not shown	Phosphate buffer (PB): Approx. round spots with a diameter of 15-100 nm and a height of 1-10 nm		Few impurities or contaminations	Few impurities or contaminations, < 2 nm width
t < 1h	Citrate buffer (CB): No fibrils PB: No fibrils	CB: Type s, curly, ring-like structures, height ~2.5 nm, width ~15 nm, length 50 – 400 nm, clumping. PB: similar as in CB.	Type s, 50-500 nm length, curly, clumping Values of height (~1.5 nm) and width (~12 nm) smaller than for Type s at pH 3.5.	Oligomer-like structures ~33 nm diameter, ~2 nm height.
Lag phase	CB: Mainly curved fibrils, Type a, b, c, and d, with branches, non-sticky. Length 1-3 µm	CB: Mainly curved, Type A, with branches, clumping. PB: Amorphous structure.	Short straight fibrils, Type b	Transition from oligomer-like structure to short straight fibrils of Type b
End of Lag phase	CB: Mainly straight fibrils Type c, d, and e. Tightly packed, don't stick, not twisted PB: Type e, 1-10 µm length, curvy, clumping	CB: Long and curved fibrils, Type c and e	Type c, and e 0.25-3.0 µm length, straight, twisted: ~0.27 µm per turn	Mainly fibrils of Type b, ~ 30-200 nm length. Very few Type c. straight.
Midpoint / t = 0 plateau	CB: t = 0 plateau: Type e, straight, not twisted, Dense mash. PB: t = 0 plateau: Type e, 2-3 µm length, curled and straight mixed, clumping.	PB: Midpoint: straight, Type c, length ~1 µm	t = 0 plateau: Type c, d, e, and f. >9 µm length, straight, twisted, no clumping, stick to each other	
t = final plateau	CB: Straight, Type e, Dense mesh. PB: Dense mesh, Type c or thicker, straight, clumping.	CB: Type c, e, and f, twisted.	Dense mesh, straight, sticky.	Type c and multiples of it. Width 45-55 nm, hight 6-10 nm, length > 9 µm, straight

Table 3.2. Description of GLP-1 aggregates imaged by AFM under different conditions.

Different peptide aggregates are defined as oligomer-like aggregates (Type o), curly and tangled fibril-like aggregates (Type s) and fibrils with a height of 0.5 - 1.5 nm (Type a), 1.5 - 2.5 nm (Type b), 2.5 - 3.5 nm (Type c), 3.5 - 5 nm (Type d), 5 - 10 nm (Type e) and >10 nm (Type f).

3.4.2 Disturbance of Aggregation Kinetics

The aim of the time-course experiments was to follow the aggregation kinetics of GLP-1 under different pH and buffer conditions. However, taking samples during a ThT assay introduces some disruption to the system which might influence the kinetics of peptide aggregation. The samples for AFM imaging were taken using a pipette which induced a gentle mixing of the samples and every time point involved an opening and closing of the plate reader. To learn more about the influence of these factors, the ThT assay data in Section 3.4.1 were analysed using a sigma function (Equation 2.5) and compared to the expected lag times of undisturbed samples, Figure 3.2A & B. The data shows that under neutral and basic conditions, the aggregation kinetics are in good agreement with the expected lag times of undisturbed samples, Figure 3.19. At pH 7.5 the lag time is most influenced by sample removal and significantly shorter, Figures 3.17A and 3.19.

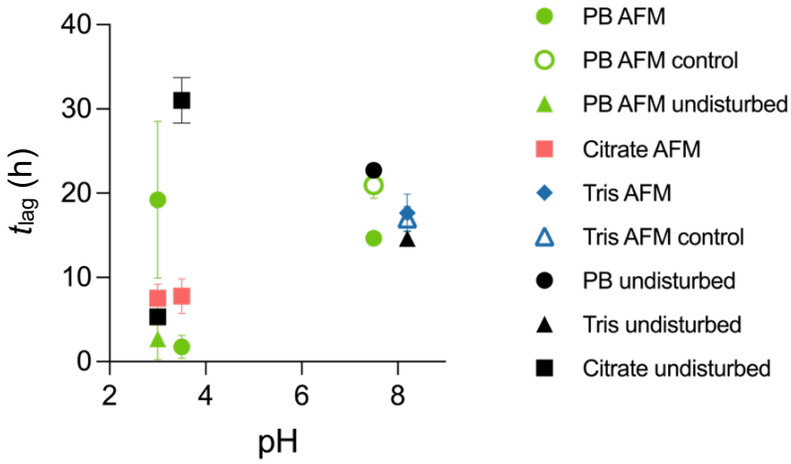


Figure 3.19. Assessment of the effect of sampling on GLP-1 aggregation kinetics.

Comparison of lag times of GLP-1 aggregation measured under different conditions. All measurements were conducted in 25 mM buffer, at 37 °C and in triplicate. Kinetics measured with disruption from sample acquisition for AFM experiments are shown as solid green circles (phosphate), solid red squares (citrate), and solid blue diamonds (Tris). Lag times calculated from a kinetic run which was completely undisrupted by sample removal and opening/closing at pH 3.0 are shown as a solid green triangle. Lag times for aggregation at pH 7.5 and 8.2, during which the plate reader was opened and closed but for which no sample was removed are shown as an open green circle (phosphate) and open blue triangle (Tris), respectively. Reference measurements (no sample removal and no opening/closing) are shown as a solid black circle (phosphate), triangle (Tris) and square (citrate). The error bars show the standard deviation of a triplicate measurement.

Under acidic conditions, the lag times of GLP-1 aggregation in phosphate buffer at pH 3.0 and 3.5, are influenced by sample removal and show different aggregation kinetics compared to the reference data, Figures 3.8A and 3.19. Interestingly, the kinetics of GLP-1 aggregation in citrate buffer at pH 3.0 shows nearly no influence of sample removal, whereas at pH 3.5 they are affected, Figure 3.19.

3.4.3 Changes in Secondary Structure During Aggregation

Far-UV CD and FT-IR time-course experiments were undertaken to elucidate the change in the secondary structure during the aggregation process. To enable the measurement of the FT-IR spectra under acidic conditions, phosphate buffer was used as citrate is known to absorb in the infrared in the same range as GLP-1. The samples for the far-UV CD and FT-IR measurements were taken at the same time and from the same plate as that used for the acquisition of samples for the AFM experiments.

The change in secondary structure was probed by far-UV CD and FT-IR at two or four different pH values, respectively, and compared with the lag time measured from the ThT assays under the same conditions, Section 3.4.2. At all the pH values studied, there is an overall change from secondary structures containing both α-helical and β-sheet structure to one that is largely β–sheet in nature at the end of the aggregation reaction as expected for the conversion of monomeric GLP-1 into amyloid fibrils.

Figure 3.20 shows the results of far-UV CD measurements at different time points during the aggregation of GLP-1 at pH 3.0 and 3.5 in 25 mM phosphate buffer. These experiments show how the main change in secondary structure of GLP-1 occurs after the lag phase and during the growth phase.

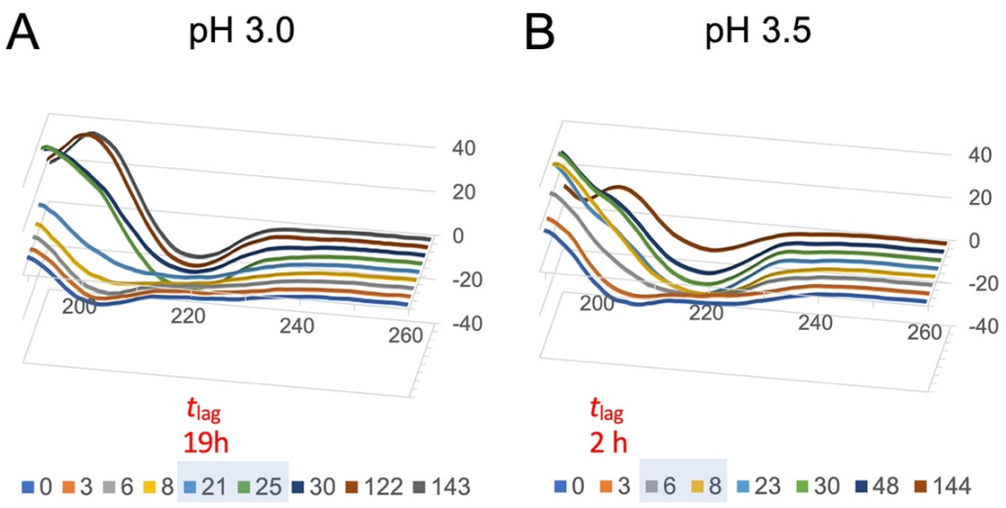


Figure 3.20. Changes in secondary structure in GLP-1 during peptide aggregation at pH 3.0 and 3.5 in phosphate buffer shown by far-UV CD.

Far-UV CD spectra of 150 µM GLP-1 measured at 25 °C. The x-axis and y-axis show the wavelength (nm) and the molar ellipticity (10^3 · deg · cm^2 · $dmol^{-1}$), respectively. The different colours represent different time points in hours. The measured lag time is shown in red. Time points showing the first major changes of the secondary structure are shaded in blue. A) 25 mM phosphate buffer pH 3.0. B) 25 mM phosphate buffer pH 3.5.

FT-IR spectra were recorded at different time points during aggregation at pH 3.5, Figure 3.21A & B, and the data indicate an even later shift from α-helical to β-sheet structure between 8 and 23 h instead of 6 to 8 h, as seen in the far-UV CD data, Figure 3.21B.

At pH 7.5, FT-IR measurements show that the major change in secondary structure occurs after the lag time and during the growth phase, Figure 3.21C. These results are in good agreement with the previously conducted far-UV CD time-course measurements, Figure 1.22C. In contrast, at pH 8.2, major changes in secondary structure occur earlier and within the lag phase, Figure 3.21D. Again, this is in good agreement with the earlier far-UV CD time-course measurements conducted by Dr Karolina Zapadka, Figure 1.22D.

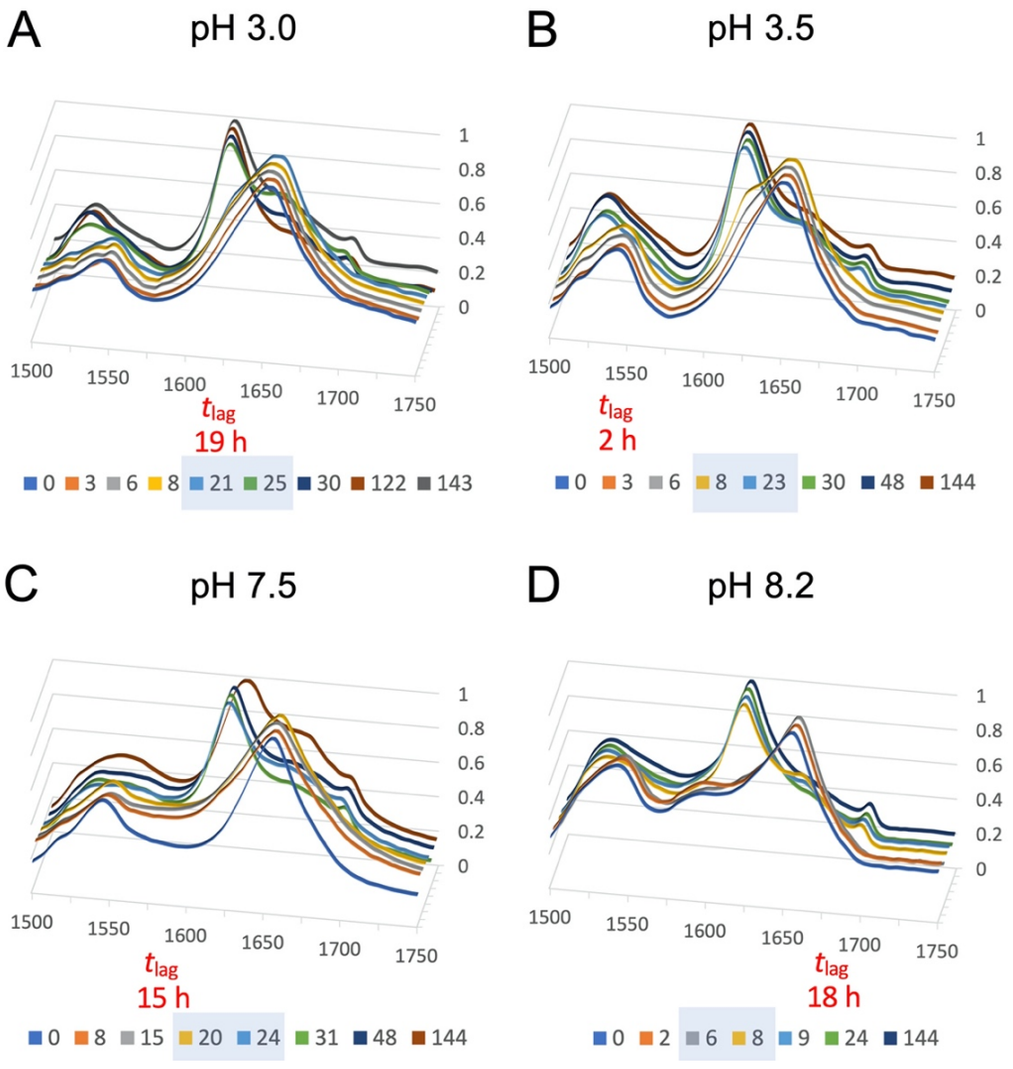


Figure 3.21. Changes in secondary structure in GLP-1 during peptide aggregation at different pHs shown by FT-IR.

FT-IR spectra of 150 µM GLP-1 measured at different time-points. The x-axis and y-axis show the wavenumber (cm^{-1}) and the absorbance (normalised), respectively. The different colours represent different time points in hours. The measured lag time is shown in red. Time points showing the first major changes of the secondary structure are shaded in blue. A) 25 mM phosphate pH 3.0. B) 25 mM phosphate pH 3.5. C) 25 mM phosphate, pH 7.5. D) 25 mM Tris, pH 8.2.

An SVD analysis of the far-UV CD and FT-IR data was performed in order to assess whether any species with secondary structure significantly different to either the monomer or fibril could be identified, and its spectrum elucidated. In all conditions, the SVD analysis demonstrates that the spectra obtained at different time points during the aggregation reaction can be deconvoluted mainly into two spectra. Typical examples of single-value deconvolution of the far-UV CD and FT-IR data are shown in Figures 3.22 and 3.23, respectively. In both cases, the data of GLP-1 at pH 3.0 at nine different time points are shown. The far-UV CD spectra in Figure 3.22A can be deconvoluted mainly into two spectra. The first, Figure 3.22B, represents the fibrillar end point, and the second represents the starting monomer, Figure 3.22C.

For the FT-IR spectra, the data can be deconvoluted mainly into two spectra, however, in this case, it is not possible to assign these two spectra to the monomer or fibrillar end point. Here, the first spectrum represents an average of both states, Figure 3.23B, and the second spectrum describes the change needed to get from the monomer to the fibrillar end point, Figure 3.23C. The fact that the first component shows an average value is most likely because the difference of the spectra is relatively symmetric between both states.

The cumulative percentages of SVD components of the far-UV CD and FT-IR spectra are shown in Figure 3.24. At pH 3.0 and 8.5, the SVD analysis of the far-UV CD data shows that the system is over 90% determined by two-states, whilst at pH 3.5 and 7.5, the values are slightly lower than this. This suggests that the system populates another state under these conditions, although not to a very high level, Figures 3.22 and 3.23D. The SVD analysis of the FT-IR data shows an even higher percentage for the sum of the first two SVD components but still lower values at pH 3.5 and 7.5 in comparison to the samples at pH 3.0 and 8.5. However, even if the system can be described by only two sigma and their vectors, this does not mean that there are no oligomeric species present, just that, if they become significantly populated that their spectra must be similar to either that of monomer or fibril.

In general, both datasets, far-UV CD and FT-IR, show high cumulative percentages already for the first sigma. The value of the first sigma of FT-IR is even higher in comparison to the first sigma of far-UV CD. This is most

likely because the FT-IR data also includes a contribution from the buffer. A subtraction of the buffer spectrum was not possible due to the drying process and unknown amounts of buffer on the observation plate in the instrument.

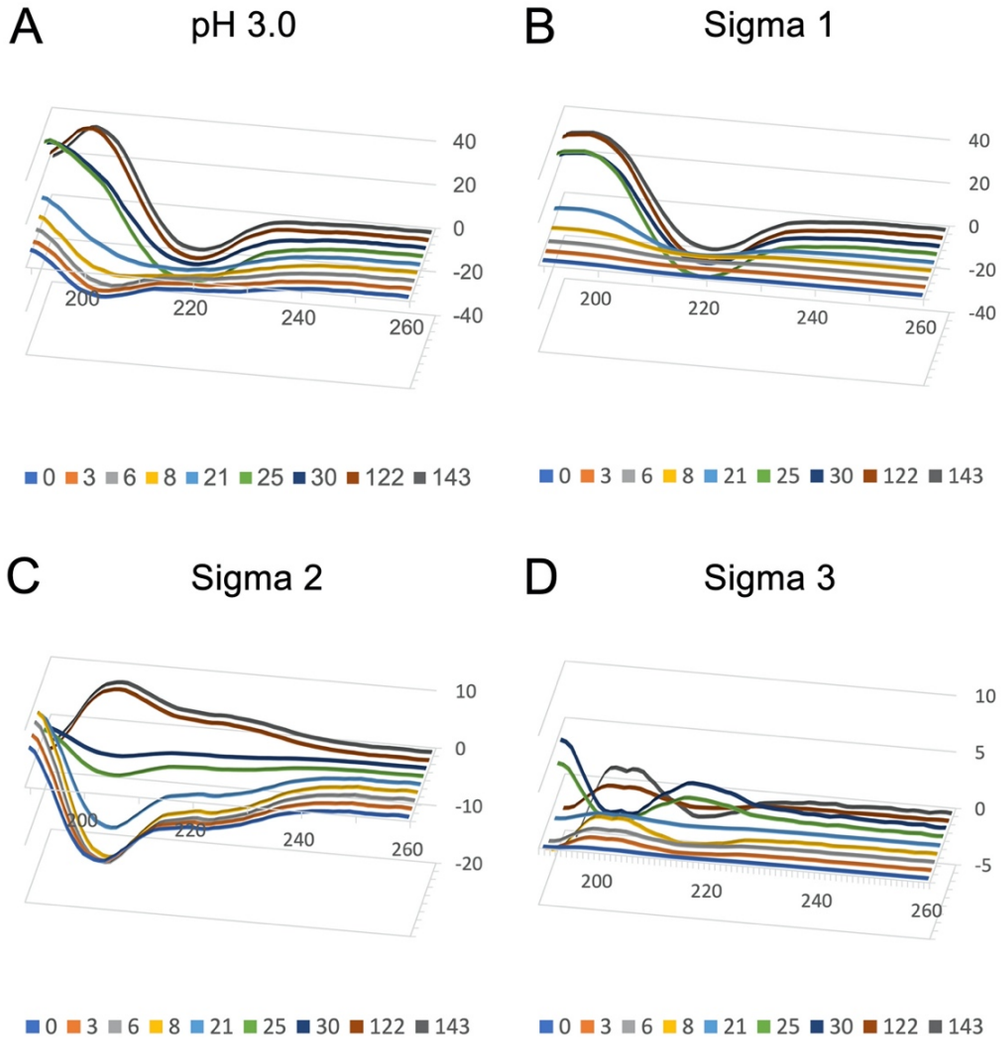


Figure 3.22. Singular value decomposition of the far-UV CD spectra of GLP-1 at pH 3.0 at different times during aggregation.

Singular value decomposition of the far-UV CD spectra of 150 µM GLP-1 in 25 mM phosphate buffer at pH 3.0, measured at nine different time points. A) Complete far-UV CD spectra. B) The first component of the analysed far-UV CD spectra. The data were calculated by multiplying $U \cdot \Sigma \cdot V^*$, with all values in Σ set to 0 except σ_1. The first component represents the far-UV CD spectra measured for aggregated GLP-1. C) The second component of the CD spectra is calculated as in B but with σ_2 as sole value instead of σ_1. The spectra represent the monomer/non-aggregated structure of GLP-1. D) Calculated third component based on σ_3. All components higher than the third show values close to 0 (data not shown).

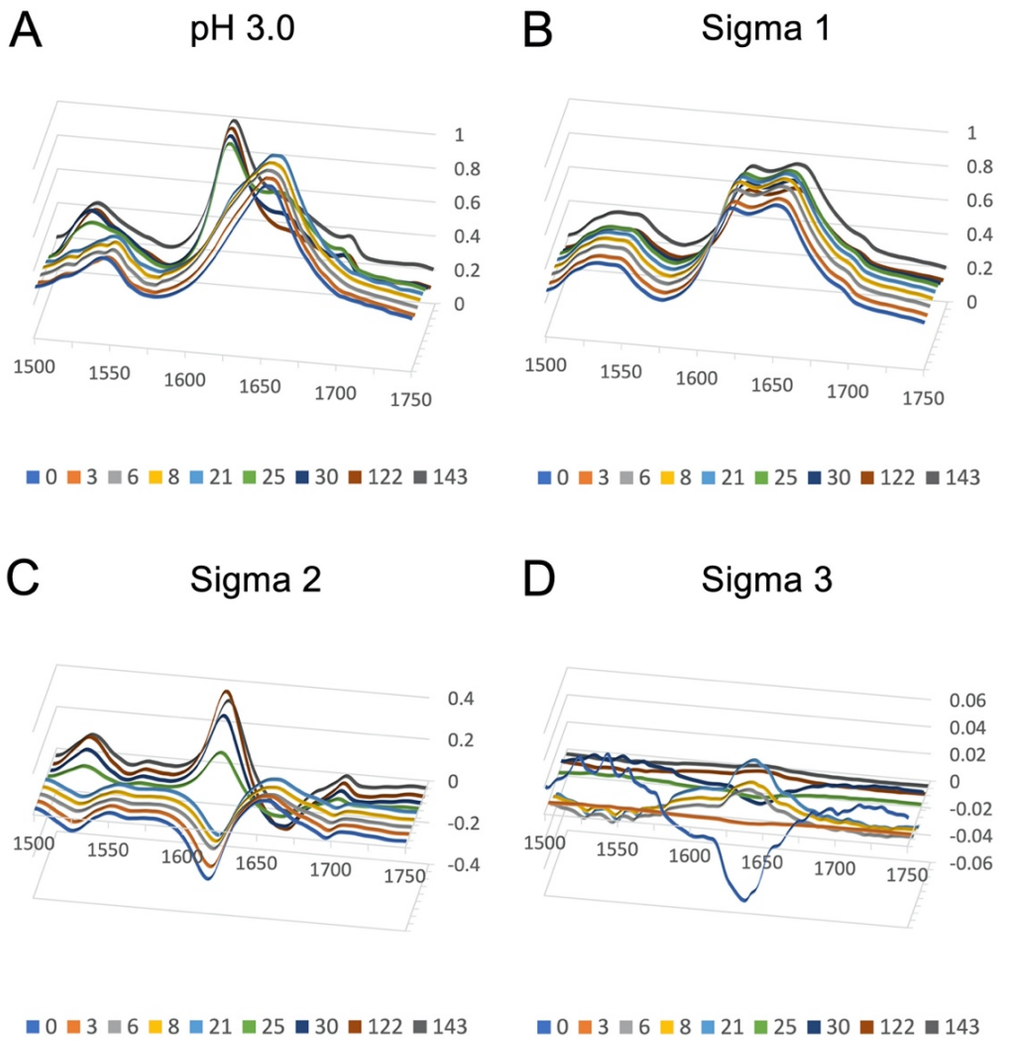


Figure 3.23. Singular value decomposition of FT-IR spectra of GLP-1 at different time points during aggregation at pH 3.0.

Singular value decomposition of FT-IR spectra of 150 µM GLP-1 in 25 mM phosphate buffer at pH 3.0, measured at nine different time points. A) Complete FT-IR spectra. B) The first component of the analysed FT-IR spectra. The data was calculated by multiplying $U \cdot \Sigma \cdot V^*$, with all values in Σ set to 0 except σ_1. The first component represents an FT-IR spectra showing average spectra for non-aggregated and aggregated GLP-1 peptide structure. C) The second component of the FT-IR spectra is calculated as in B but with σ_2 as sole value instead of σ_1. The spectra represent the change between monomer/non-aggregated and aggregated structure of GLP-1. D) Calculated third component based on σ_3. All components higher than the third show values close to 0 (data not shown).

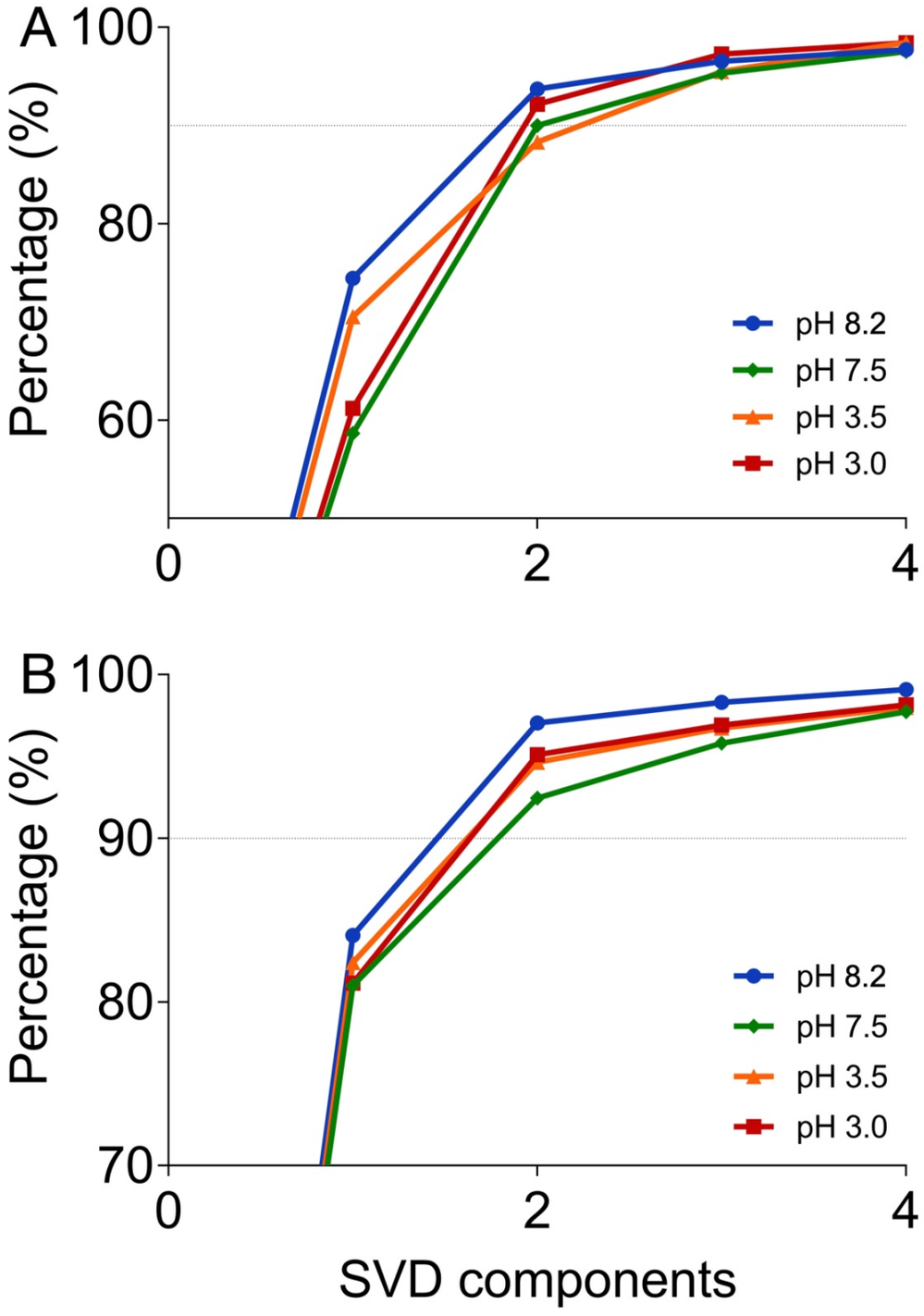


Figure 3.24. Cumulative percentage of single-value decomposition (SVD) components.

A) Cumulative percentage of SVD components of far-UV CD spectra. 150 µM GLP-1 in 25 mM buffer measured at 25°C. The measured samples were aggregated in phosphate buffer at pH 3.0, phosphate buffer at pH 3.5, phosphate buffer at pH 7.5 and Tris buffer at pH 8.5. B) Cumulative percentage of SVD components of FT-IR spectra. 150 µM GLP-1 in 25 mM buffer measured at 25 °C. The measured samples were aggregated in phosphate buffer at pH 3.0, phosphate buffer at pH 3.5, phosphate buffer at pH 7.5 and Tris buffer at pH 8.5.

3.4.4 Size of GLP-1 Aggregates in Solution

To elucidate the size of the aggregates formed during GLP-1 aggregation, DLS measurements were performed at the same time points as the AFM, far-UV CD and FT-IR data. The conducted measurements at pH 3.0, 3.5 and 6.5 were used together with previous data obtained for GLP-1 at pH 7.5, 7.7 and 8.2[5] to establish how the size of species varies over time over a wide range of pH values. It is essential to mention that the data are presented as both intensity and number distributions. The intensity measurements have a stronger signal for larger particle sizes so tend to be dominated by these species, however, this measurement does not depend upon any model used. N.B., the Zetasizer software calculates the number distribution based on the intensity distribution using assumptions about the samples' shape, density, and optical properties. Throughout this section measurements given are for the diameter of species present.

Freshly prepared samples of GLP-1 at pH 3.0, show a peak corresponding to monomer and possibly small oligomers at 5-6 nm, in addition to a peak indicting species with a size of approx.1000-2000 nm caused by a small number of particles which scatter the light with high intensity. Still, the smaller molecular weight peak represents only 5-10% of the sample concentration and disappears completely after the first measurement, presumably as monomers and small oligomers are converted into larger aggregates and ultimately fibrils. In general, there are large differences between the single measurements at different time points, Figure 3.25A & B. This is most likely caused by ongoing aggregation reactions of the samples which are not at equilibrium or simply the settlement of the aggregates at the bottom of the cuvette during the measurement. To be consistent with the previous experiments, 150 μM GLP-1 was used, and it should be noted that this is lower than the optimal concentration needed for DLS measurements and this may also contribute to the variation in signal between time points.

At pH 3.5, the first peak of the earliest time points shows even smaller particles around 3-4 nm, which corresponds with approximately 5% of the total peptide concentration. Additionally, larger particles are also present. They have a diameter of about 50-200 nm, tenfold smaller than the larger aggregates

observed at pH 3.0. After 3, 6 and 8 h smaller aggregates can be observed, with diameters of approximately 10 - 100 nm, Figure 3.25C & D.

At pH 6.5, two measurements were conducted. Immediately after the peptide was dissolved and filtered, the intensity data shows species and aggregates around 3-4 and 50-200 nm. After the ThT assay (144 h), during which no increase of ThT fluorescence was observed, the monomer/dimer peaks vanish and the number of larger aggregates between 30 and 700 nm increases. Additionally, a few aggregates with diameters larger than 2000 nm are present, Figure 3.25E. At time point zero the majority of the aggregates are between 3-4 nm and after the ThT assay between 20-80 nm, Figure 3.25E.

In general, it is assumed that the smaller molecular weight species between 2 – 6 nm are most likely monomer or dimer, and this is only observed for the first measurement of the freshly dissolved peptide solution. This is similar to the findings for GLP-1 at pH 7.5, 7.7 and 8.2.[5] These results suggest that the monomer/dimer rapidly converts into other species.

At pH 3.5, the data at time points within the lag phase, show large numbers of oligomeric-like aggregates around 10 – 100 nm. Similarly sized oligomeric-like aggregates (20 – 100 nm) were found at pH 6.5 at the end of the ThT assay. As the results of the AFM and ThT assays showed no evidence for fibrillar species, this suggests that the oligomeric-species observed at both pH 3.5 and 6.5 are related and likely to be off-pathway, not converting directly into amyloid fibrils. At pH 6.5, both these species are highly stable and do not convert back into monomer, on-pathway species or fibrils. Alternatively, fibrils may not be stable under these conditions.

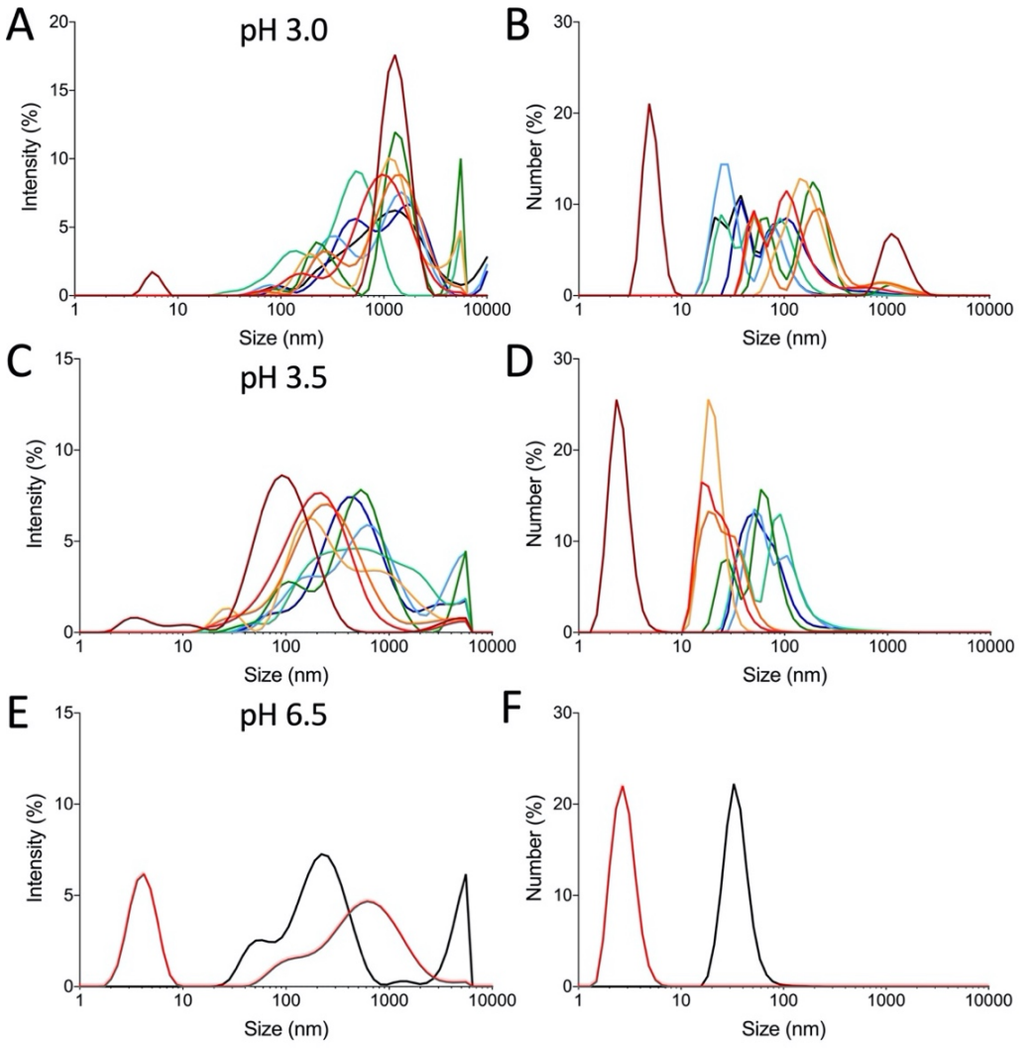


Figure 3.25. Results of DLS measurements to determine the size of GLP-1 species in solution at different time points during aggregation.

DLS measurements of 150 µM GLP-1 A) Intensity vs. size of particles (diameter) in 25 mM phosphate buffer at pH 3.0. B) Number of particles vs. size of particles (diameter) in 25 mM phosphate buffer at pH 3.0. A & B) Time points taken after 0 h (dark red), 3 h (red), 6 h (dark orange), 8 h (orange), 21 h (dark green), 25 h (bright green), 30 h (bright blue), 122 h (blue) and 143 h (dark blue). C) Intensity vs. size of particles (diameter) in 25 mM phosphate buffer at pH 3.5. D) Number of particles vs. size of particles (diameter) in 25 mM phosphate buffer at pH 3.5. C & D) Time points taken after 0 h (dark red), 3 h (red), 6 h (dark orange), 8 h (orange), 23 h (dark green), 30 h (bright green), 48 h (bright blue) and 144 h (blue). E) Intensity vs. size of particles (diameter) in 25 mM citrate buffer at pH 6.5. F) Number of particles vs. size of particles (diameter) in 25 mM citrate buffer at pH 6.5. E & F) Time points taken after 0 h (red) and 141 h (black).

3.5 Discussion

Problems associated with the physical stability of many potentially otherwise very efficacious therapeutic peptides has hindered or prevented their use in the clinic. Amongst other forms of physical instability, many peptides are known to form amyloid-like fibrils at their high formulation concentrations. In order to be able to better develop peptide-based drugs with improved physical stability profiles, a better understanding of the underlying mechanisms of aggregation and amyloid formation are required. The therapeutic peptide GLP-1 has been shown to form amyloid fibrils and this has been studied in detail at pH 7.5-8.2.[5] An unusual change in the aggregation kinetics between pH 7.5 and 8.2 has been attributed to the population of off-pathway oligomers at pH 7.5, in addition to the on-pathway oligomers which convert into the amyloid fibrils. There is increasing evidence for, and awareness of, off-pathway species,[5,89,130,343–349] however, few detailed studies of such species and their impact on aggregation kinetics have yet to be published.

Following on from the work on GLP-1 at pH 7.5 - 8.2, here an investigation of the physical stability of GLP-1 at basic, neutral and acidic pH values is reported. The focus was on time-course measurements during which a number of biophysical techniques were employed that report on the structure and size of species in solution at various times during the aggregation reaction. The objectives of the study were two-fold, i) to understand in more detail the conditions in which off-pathway species become significantly populated in order for us to better understand the nature of these species, and ii) to gain further information on species present in solution at different time points under conditions which favour the population of on- and off-pathway oligomers, particularly focussing on processes that occur within the lag time.

3.5.1 Physical Evidence for Off-Pathway Oligomeric Aggregates

The unusual concentration dependence of GLP-1 aggregation kinetics at pH 7.5 has been previously explained by the influence of off-pathway oligomeric aggregates.[5] The experiments in this Chapter were conducted to understand

why such an unusual behaviour can be found not only at pH 7.5 but at pH 3.5 suggesting that off-pathway oligomers are also populated under acidic conditions at pH 3.5, but interestingly not at pH 3.0. For both pH 3.5 and 7.5, the study also reveals further evidence for off-pathway species. For example, analysis of the far-UV CD and FT-IR data with SVD showed that an additional species with characteristics different to either the monomer and fibrils is more evident and populated at pH 3.5 and 7.5 in comparison to pH 3.0 and 8.2. In addition, and based on the results of the AFM experiments, results suggest that such a putative off-pathway oligomer may be the short and curly fibril-like aggregates observed at early time points at some pH values. These species appear to have hydrodynamic radii of about 20-80 nm as observed by DLS at pH 3.5 and 6.5.

3.5.2 Characteristics of GLP-1 Aggregates and Fibrils Formed Under Acidic and Basic Conditions.

As mentioned in the previous section, short and curly fibril-like species have been seen under both conditions known to show off-pathway behaviour. Under acidic conditions, the curly fibril-like aggregates are thicker in comparison to the thinnest types of fibrils observed in the same sample, consistent with the idea that these aggregates are off-pathway because they would have to undergo dissociation and a significant change in structure to transform them into the thinnest types of fibrils observed.

While showing the same characteristics as the short curly fibrils formed under acidic conditions, the dimensions of the curly fibril-like aggregates imaged at neutral pH are thinner than the curly fibril-like aggregates imaged at pH 3.5. Their dimensions are similar to the height and width of the observed thin fibrils at pH 7.5. Although these aggregates are similar in width to other fibrils formed, they are still likely to be off-pathway, as again they would need to undergo a structural transformation in order to form the more rigid structure adopted by fibrils.

The aggregation of GLP-1 under basic conditions (pH 8.2) follows kinetics typical of a nucleation-elongation mechanism. The AFM data suggests that the nucleation may occur in small on-pathway, oligomer-like aggregates with a

diameter slightly bigger than the width of the fibril formed from the nucleus.

In general, during the lag phase and partly at the beginning of the growth phase, several different sized fibrils have been detected, for example, at pH 3.0. Their dimensions suggest that fibrils are multiples of the thinnest fibril type, Type a, which is possibly a protofibril. Fibrils of Type b are most likely formed from two Type a species. The number of protofibrils in fibrils of Type b, c and d is estimated to be 2, 4, and 8, respectively. However, it cannot be excluded that some fibrils are formed from different numbers of protofibrils. As fibrils of Type c and e appear under most or even all pH conditions, it is reasonable to speculate that all fibrillar species share the same overall assembly pattern.

3.5.3 Challenges and Further Results from the Time-Course Measurements

The time-course measurements lead to important insights regarding the nature of the aggregates formed by GLP-1 and the changes in secondary structure during the aggregation process. However, especially under acidic conditions, it is evident that the removal of samples for the AFM and other measurements has an influence on the aggregation. Additionally, it cannot be excluded that the phosphate buffer shows different behaviour than the citrate buffer. While the aggregates formed in citrate or phosphate buffer were similar, the measured lag times showed some differences. These all make it more challenging to draw conclusions from the data collected under acidic conditions.

The data recorded under neutral and basic conditions is, in contrast, more reliable. At pH 7.5, the secondary structure stays in the α-helical conformation during the whole lag phase, suggesting that off-pathway oligomers formed may have α-helical structure. In the basic regime, the secondary structure changes early from α-helical to β-sheet, this transformation happens even before the lag time and before any significant ThT signal was observed. These results suggest that an intermediate, non-fibrillar species which has considerable β-sheet content but which cannot bind to ThT is populated.

3.5.4 The Relationship Between Aggregation Propensity and pI and Net Charge

The aggregation kinetics show that the lag time for fibril formation at pH 3.0-3.5 varies between 5 and 30 h and the half times vary from 10-40 h, depending upon pH and peptide concentration, Figures 1.17 and 1.18. These values are comparable to those measured at pH 7.5-8.2. The approx. net charge GLP-1 carries under these conditions is +3 to +4 at low pH, compared to -1 to -2 under slightly basic conditions. Despite having higher net charge values under acidic conditions, collectively, the results of the aggregation assays conducted at different pH values show that there is some symmetry in the behaviour, i.e., aggregation propensity, lag time and peptide concentration dependence, around the pI value of GLP-1. The simplest explanation of this phenomenon is that off-pathway oligomers which affect the kinetics are populated close to the pI, i.e., under conditions where the net charge on the peptide is close to zero. It is interesting to note that at pH 6.5, the pH closest to the pI, GLP-1 appears to populate off-pathway species that do not convert into fibrils as no fibrils were detected either by ThT fluorescence or in AFM imaging.

3.5.5 Conclusions

In summary, the extensive study of GLP-1 presented here has led to a more detailed understanding of the aggregation kinetics of GLP-1 under several conditions. Starting from the observation that the unusual peptide concentration dependence of the aggregation kinetics is not restricted to neutral and basic conditions but can also be found under acidic conditions, it can be concluded that off-pathway oligomer formation occurs at pH 3.5, 6.5 and 7.5. Time-course studies at these pH values reveal the possible structure of these oligomers. As there were no observed changes in secondary structure during the lag phase and before there is any significant ThT fluorescence under these conditions, this suggests that this type of oligomer has α-helical structure. The pH range over which there is evidence of off-pathway oligomer formation suggests that off-pathway oligomers are populated under conditions close to the pI of the peptide and therefore that net charge plays a significant role in determining the species

formed (on *versus* off-pathway oligomers) and therefore the rate of aggregation.

4 The Influence of C-terminal Amidation on the Physical Stability of GLP-1

4.1 Introduction

The physical stability of therapeutic peptides and proteins is of great importance as their physical instability, i.e., their tendency to self-associate into aggregated states, is a significant problem for the pharmaceutical industry. In previous work, it was shown that the kinetics of aggregation of GLP-1 depends critically on pH and that a switch in aggregation mechanism takes place over a narrow pH range near pH 3 to 3.5 and pH 7.5 to 8. This has been attributed to a change in mechanism from one in which on-pathway oligomers dominate, to one in which off-pathway oligomeric species become populated and drastically affect the kinetic behaviour.

In this Chapter, a detailed study on the effect of C-terminal amidation on the physical stability of GLP-1 is undertaken. The purpose of this study is to assess the influence of chemical modification, in this case the amidation of the C-terminus of a therapeutic peptide, on its physical stability. In comparison to other modifications, amidation represents a rather small change to the peptide and is often used in combination with other modifications. It is also found in

naturally occurring peptides and, for some neuropeptides, amidation is crucial for their full biological activity.[350]

Here, the physical stability of GLP-1 and GLP-1 amidated at the C-terminus (Am-GLP-1) are compared and the differences and similarities in their physical stability highlighted. The focus is on their behaviour at different pH values and the peptide concentration dependence of their aggregation kinetics. Numerous biophysical measurements including ThT assay, far-UV CD, SEM, ANS and tryptophan fluorescence were used to gain insight into the mechanism of aggregation into amyloid fibrils and the nature of species populated during fibrillation under different conditions.

RESULTS

4.2 Molecular Weight and Purity of GLP-1 and Am-GLP-1

The molecular weight and purity of GLP-1 and Am-GLP-1 were determined using LC-MS measurements. The UV absorbance at 210 nm from the liquid-chromatography step shows a single narrow peak for both peptides, Figures 4.1A and 4.2A, indicating the samples have high purity. The same peak as observed with UV absorbance was observed in the total ion current (TIC) data, Figures 4.1B and 4.2B. However, both peptides also show additional smaller peaks at longer elution times. These observed ions do not absorb in the UV and are most likely hydrophobic in nature, hence the longer retention time. At the current time, the identity of these smaller peaks is not known, however, it is currently under investigation as part of a new PhD project within the Jackson group.

From the mass spectrometry data, it is possible to calculate the molecular weight of the peptide from the measured m/z of the protonated peptide ions. The data shows that GLP-1 is mainly present as a +3 molecular ion (1119.42 m/z). Minor peaks can also be seen for the +4 molecular ion (839.72 m/z), and +2 molecular ion (1678.91 m/z). From these data, the determined molecular weight of GLP-1 is 3355.3 ± 0.5 Da which is in agreement with the theoretical value of 3355.66 Da as well as the molecular mass of 3355.71 Da measured using electrospray ionisation-mass spectrometry reported by Bachem, Figure 4.1.

The mass spectrum of Am-GLP-1 is similarly to GLP-1 mainly present as a +3 molecular ion (1119.1 m/z), Figure 4.2. The other peaks represent the +5 molecular ion (671.87 m/z), +4 molecular ion (839.55 m/z), and +2 molecular ion (1678.43 m/z). The molecular weight determined is 3354.68 ± 0.5 Da which agrees with the theoretical value of 3354.68 Da as well as the relative molecular mass of 3354.93 Da obtained using electrospray ionisation-mass spectrometry by Bachem.

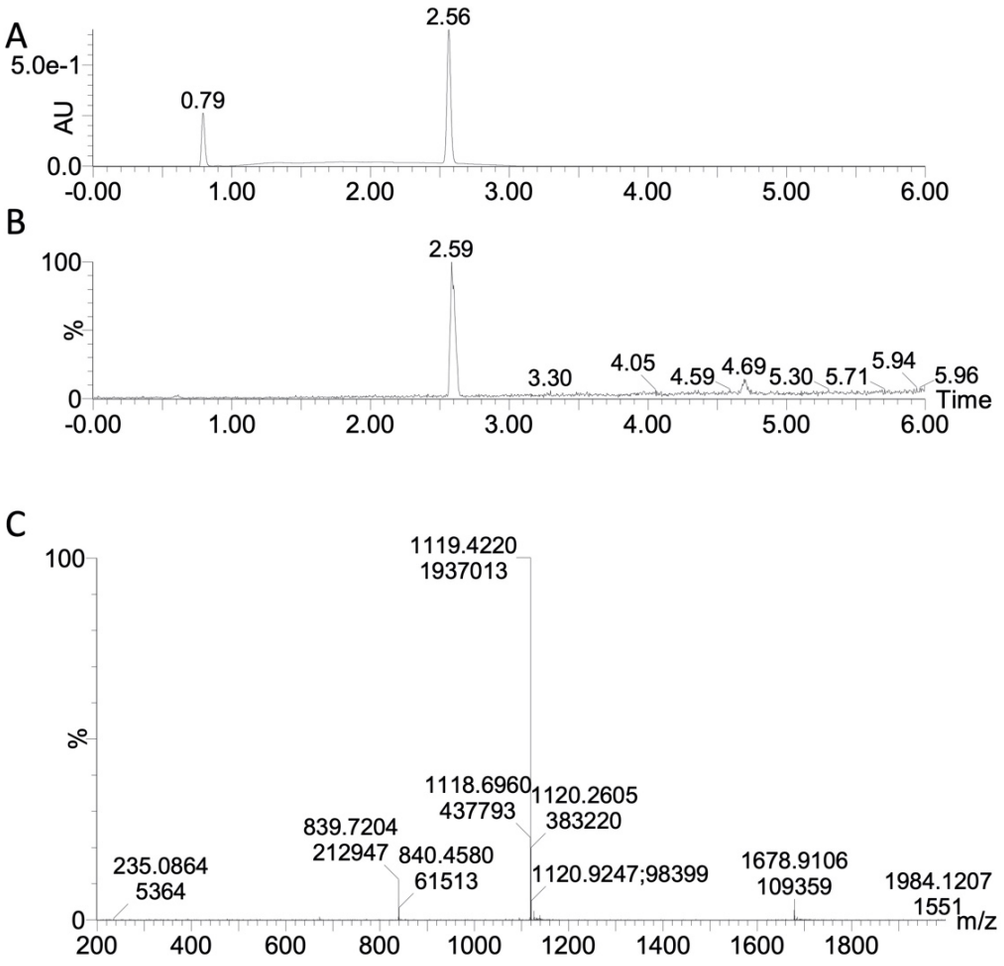


Figure 4.1. LC-MS of GLP-1.

A) UV absorbance signal at 210 nm. The peak at 0.79 min is caused by the injection of the sample. It is influenced by the different percentage of acetonitrile in the sample in comparison to the LC gradient as well as the presence of salt in the sample. The x-axis is given in minutes. B) Total ion current (TIC) chromatogram. The x-axis is given in minutes. C) Mass spectrum of peak eluting at 2.59 min.

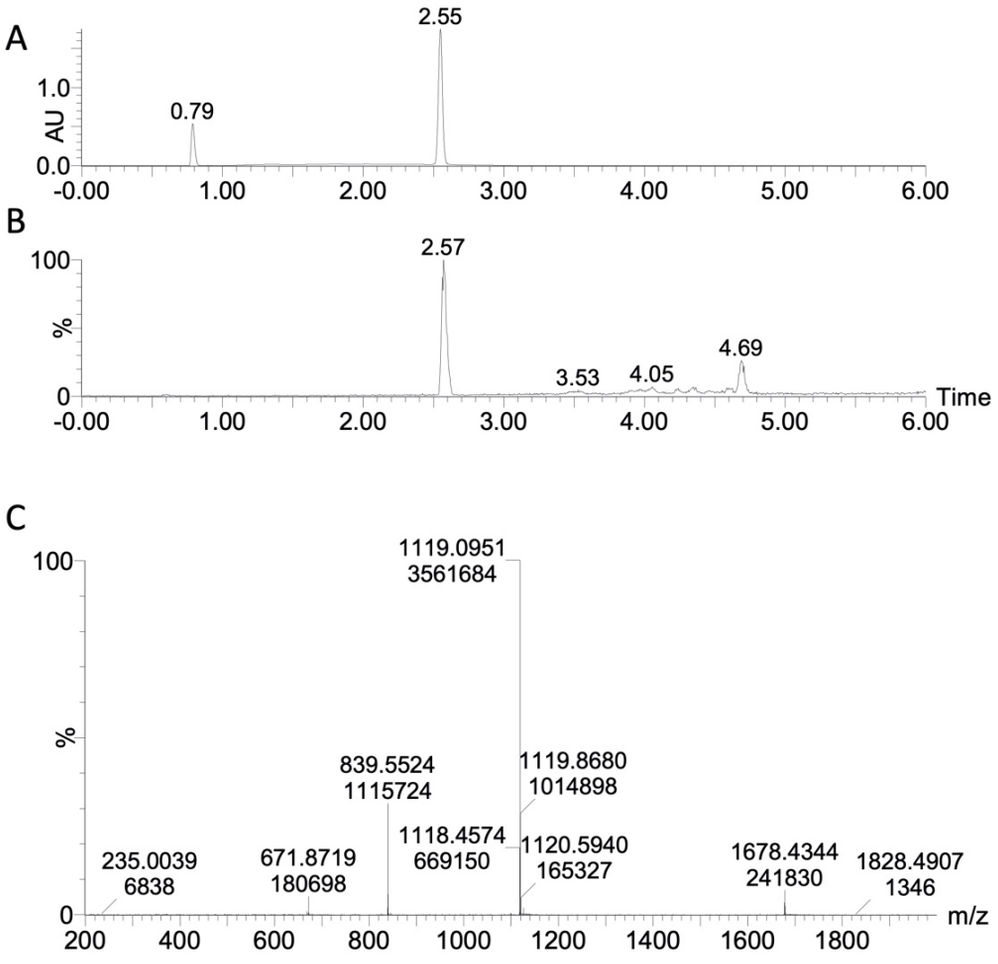


Figure 4.2. LC-MS of Am-GLP-1.

A) UV absorbance signal at 210 nm. The peak at 0.79 min is caused by the injection of the sample. It is influenced by the different percentage of acetonitrile in the sample in comparison to the LC gradient as well as the presence of salt in the sample. The x-axis is given in minutes. B) Total ion current (TIC) chromatogram. The x-axis is given in minutes. C) Mass spectrum of peak eluting at 2.57 min.

4.3 The Influence of C-terminal Amidation on Peptide Solubility

The solubility of Am-GLP-1 was measured and compared to that of GLP-1, and significant differences were observed. GLP-1 shows a minimum in solubility around pH 4.5 whilst for Am-GLP-1 it is around pH 7.5, Figure 4.3A. The observed minimum in solubility for Am-GLP-1 is much broader than the one observed for GLP-1 and stretches from pH 5.0 to 8.5 in comparison to pH 4.0 to 5.5 for GLP-1, Figure 4.3A. These differences are likely to be due to the difference in net charge as amidation results in the removal of the negative charge at the carboxy terminus at pH values above the pK_a (2.3) of the C-terminal glycine carboxyl group. Figure 4.3B shows the net charge of GLP-1 and Am-GLP-1 at different pH values calculated using the Henderson-Hasselbalch equation and standard pK_a and pK_b values for N- and C-terminus and ionisable side chains. The details of the calculation are shown in Section 2.15 and Table 2.2. Ignoring other influences, such as changes in conformation, amidation of the C-terminus leads to a steady net charge difference of 1 between GLP-1 and Am-GLP-1 at most pH values until below pH 2.3 when the C-terminus of GLP-1 becomes protonated and the two peptides have the same net charge. It is also clear that the net charge curve for the two peptides cuts the zero net charge line at different positions and with different slopes, Figure 4.3B. It means that the data predicts the net charge of GLP-1 to be within a range between ± 0.5 from approx. pH 5.0 to 6.0 in comparison to the wider range of pH 6.0 to 9.0 in the case of Am-GLP-1. This correlates well with the solubility data for GLP-1 and Am-GLP-1 shown in Figure 4.3A.

The pI values of GLP-1 and Am-GLP-1, calculated theoretically using standard pK_a values, are 5.4 and 7.6, respectively and can be compared to the values determined experimentally using imaged capillary isoelectric focussing (iCIEF) shown in Figure 4.3C. For GLP-1, a pI of 5.9 was determined, which is in good agreement to the theoretical value of 5.4. In the case of Am-GLP-1, the difference between the calculated pI of 7.6 and the measured value of 6.8 is slightly larger. This difference may be due to the fact that Am-GLP-1 has a

very broad minimum of net charge around the pI in comparison to GLP-1. Small alterations of the net charge, for example caused by changes in conformation, may therefore have a bigger influence on the pI.

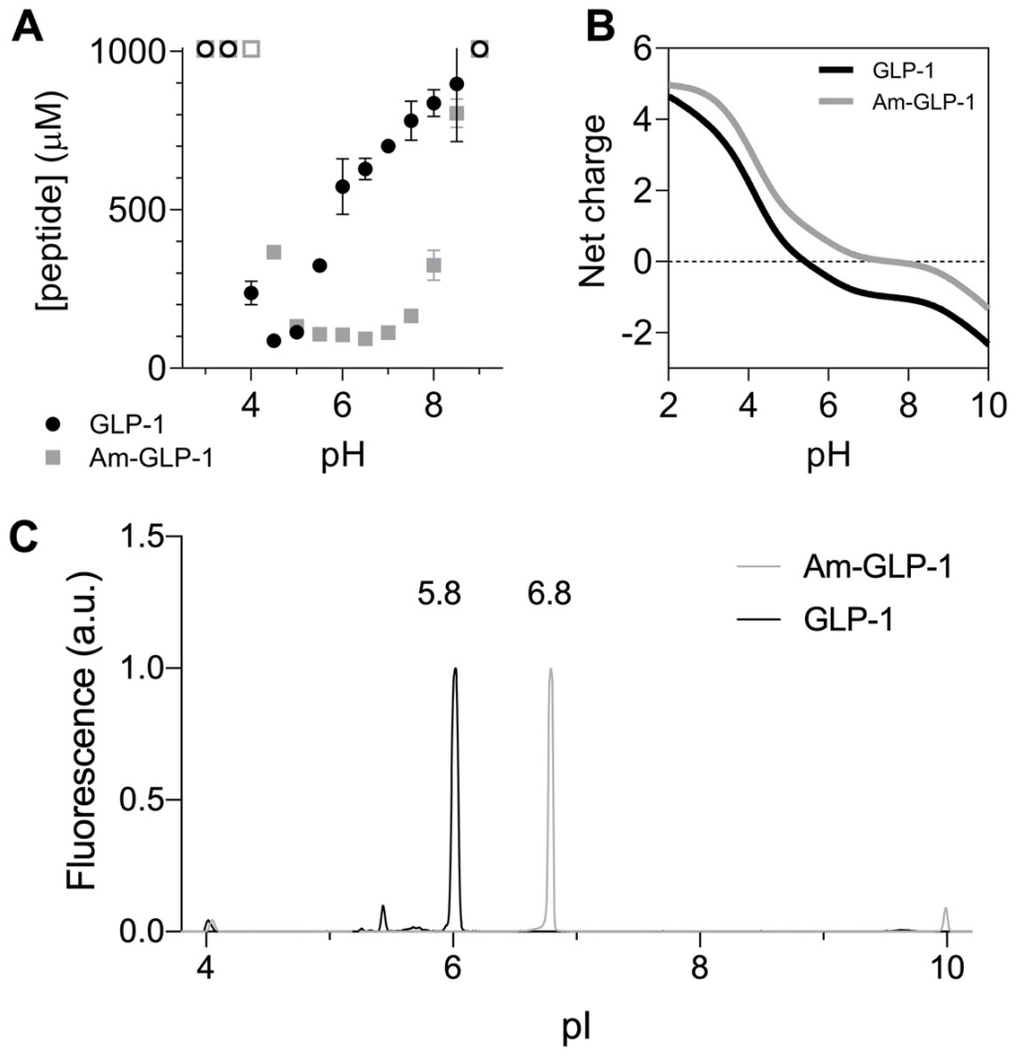


Figure 4.3. Solubility, net charge of Am-GLP-1 and GLP-1.

A) Peptide concentration measured after dissolving approximately 0.2 mg in 50 μL of buffer: GLP-1 (black closed circles) or Am-GLP-1 (grey closed squares). The open circles indicate a concentration of at least 1000 μM. The error bars show the standard deviation of three peptide concentration measurements. B) Calculated net charge *versus* pH based on the peptide amino-acid sequence and standard pK_a values. Note the rather flat intersection of the net-charge curve of Am-GLP-1 and the zero value. C) pI measurement for Am-GLP-1 and GLP-1 using imaged capillary isoelectric focussing.

4.4 The Peptides Have a Lower Aggregation Propensity Close to Their pI Values

The aggregation kinetics of the GLP-1 peptides were measured using a ThT assay. In Figure 4.4A & B, two three-dimensional aggregation landscapes are shown which are composed of ThT data measured over time at different pH values between pH 3.0 and 8.7 and where the peptide concentration was kept constant at 75 μM. These data allow one to quickly visualise the aggregation propensity of GLP-1 and Am-GLP-1 over a wide range of pH values giving a broad overview of the influence of pH on the aggregation behaviour of both peptides. There is a strong correlation between the net charge and solubility of the peptides and their propensity to aggregate. GLP-1 shows a minimum in aggregation around its pI and, as net charge and solubility increases between pH 5.5 and 7.5, this results in a shortening of the lag times. Am-GLP-1 also shows a similar relationship between net charge, solubility and aggregation propensity and has a minimum of aggregation between pH 6.0 and 8.2, which corresponds to the range in which the net charge on the peptide is between +0.5 and -0.5.

Starting from these observations, four different pH values were chosen at which the aggregation of Am-GLP-1 was investigated in further detail. In particular, the effect of peptide concentration on the aggregation kinetics was measured thereby enabling us to compare it with our previously published data on GLP-1.[5]

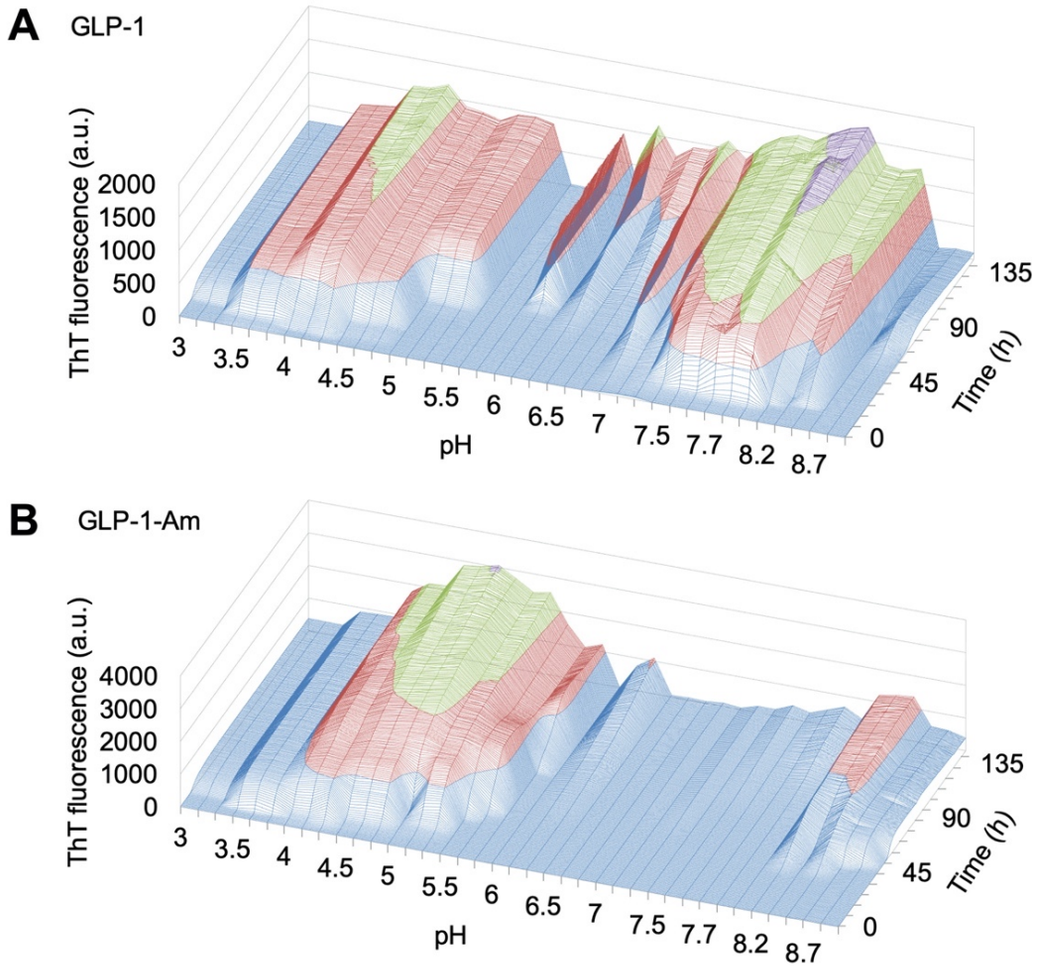


Figure 4.4. Aggregation landscapes for GLP-1 and Am-GLP-1.

A) ThT assays of peptide aggregation for GLP-1 at 75 µM from pH 3.0 to 8.7 at 37 °C. The buffers are citrate (pH 3.0 - 6.0), phosphate (pH 6.5 - 7.5) and Tris (pH 7.7 – 8.7, 37 °C) with a buffer concentration of 25 mM. For each pH the results of triplicate runs are shown. The x-axis shows the pH value, z-axis time (h) and the y-axis ThT fluorescence (a.u.) where the signal between 0-500 is blue, 500-1000 (red), 1000-1500 (green) and 1500-2000 (purple). B) ThT fluorescence for Am-GLP-1 at 75 µM from pH 3.0 to 8.7 at 37 °C. For each pH the results of triplicate runs are shown. The axes are the same as in Figure 4.3C.

4.5 Aggregation Kinetics of Am-GLP-1: Peptide Concentration Dependence

To learn more about the influence of the amidation of the C-terminus on the physical stability of GLP-1, kinetic aggregation assays were performed for Am-GLP-1 at different peptide concentrations. The aggregation kinetics were monitored using ThT fluorescence and all measurements were run in triplicate

and the experiments repeated at least two times on different days. The pH values were chosen taking into account the previous ThT fluorescence measurements on Am-GLP-1 (Figure 4.3), as well as the previous measurements conducted on GLP-1 at pH 3.0, 3.5, 7.5 and 8.2, Figures 3.1 and 1.17.[5,351]

Four pH values, 3.0, 4.0, 8.2 and 8.7, were chosen and measurements were performed at six different peptide concentrations between 25-150 μM, Figure 4.5A-D. In many cases, it was possible to describe and fit the ThT fluorescence data to a sigmoidal curve with a flat lag phase, a steep growth phase and a plateau. The kinetic parameters $t_{1/2}$ and t_{lag} were determined from a standard sigmoidal fit of the data to Equations 2.4 and 2.5, respectively.

Figure 4.5A shows the ThT aggregation data for Am-GLP-1 in 25 mM citrate buffer at pH 3.0 and both the lag time and $t_{1/2}$ decrease with increasing peptide concentration (Figure 4.6A & C) typical of a nucleation-polymerisation mechanism with predominantly on-pathway oligomers. It is interesting to note that at low peptide concentrations the data fit well to a sigmoidal function, however, at higher peptide concentrations the data fit less well.

Data for Am-GLP-1 aggregation in citrate buffer at pH 4.0 are shown in Figure 4.5B. Under these conditions, the lag time and $t_{1/2}$ are independent of the peptide concentration, in contrast to the results obtained at pH 3.0. It is possible to explain these results by the population of off-pathway oligomeric species under these conditions,[87] as has been shown for GLP-1 at pH 7.5.[5] At pH 4.0, the ThT data shows some deviation from the standard sigmoidal behaviour. First, at low peptide concentrations, the growth phase of Am-GLP-1 shows two distinct regions – an initial very steep slope which changes to a slope that is less steep before the plateau region is reached, Figure 4.5E. In addition, at higher peptide concentrations, a linear increase in the baseline within the lag phase was observed before it changes over to the typical sigmoidal growth phase, Figure 4.5F. At intermediate peptide concentrations the aggregation can alternate between the two types of behaviour leading to a higher variance between runs, Figure 4.5B, than seen at pH 3.0.

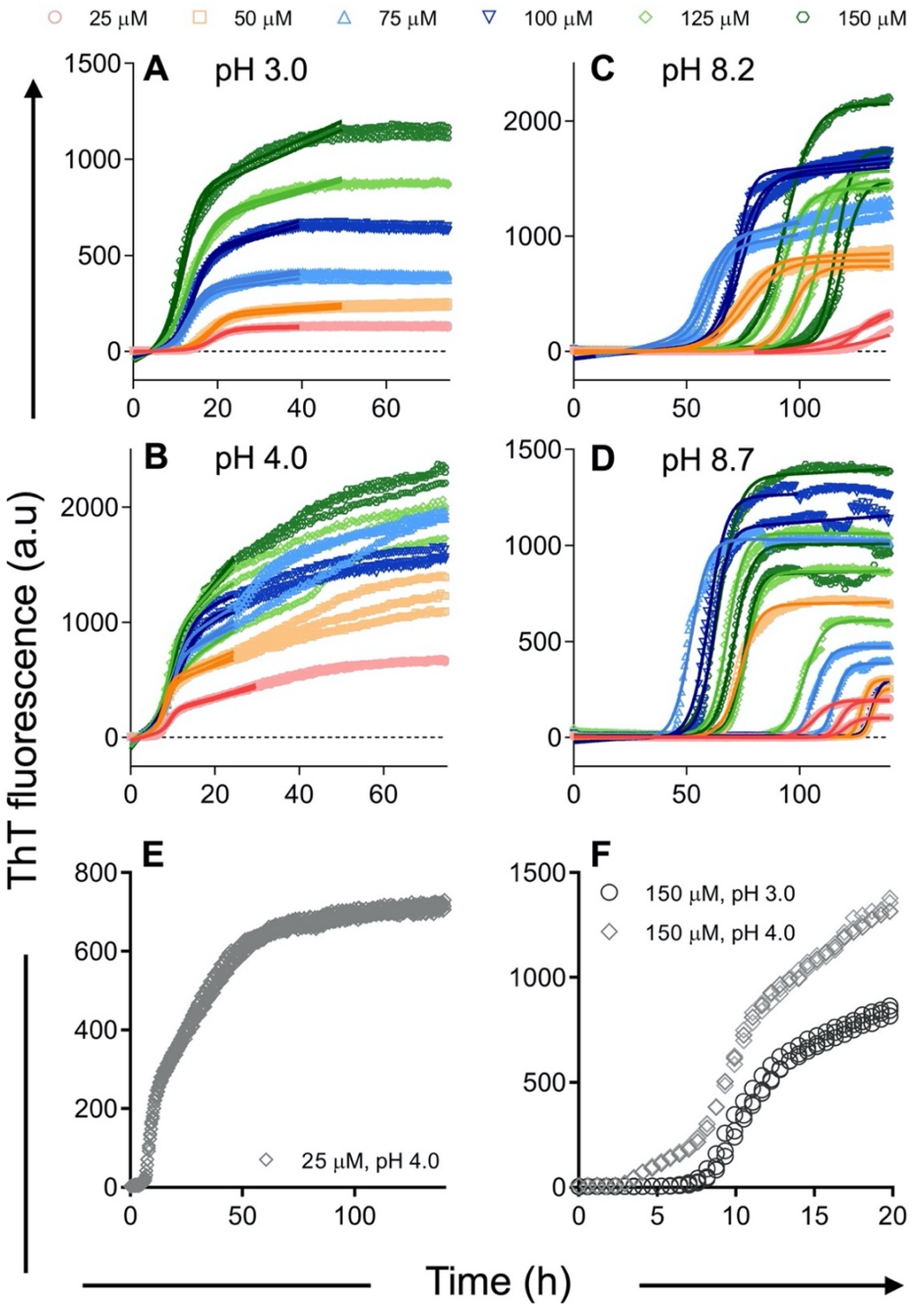


Figure 4.5. ThT aggregation assays of Am-GLP-1 at different peptide concentrations and pH values.

All samples were measured at the following peptide concentrations: 25 μM (pink), 50 μM (orange), 75 μM (blue), 100 μM (dark blue), 125 μM (green), and 150 μM (dark green). The data were fitted with Equation 2.4, the best fits are shown by the solid lines. All experiments performed at 37 °C. A) Am-GLP-1 in 25 mM citrate buffer at pH 3.0 B) Am-GLP-1 in 25 mM citrate buffer at pH 4.0. C) Am-GLP-1 in 25 mM Tris buffer at pH 8.2. D) Am-GLP-1 in 25 mM Tris buffer at pH 8.7. E) 25 μM Am-GLP-1 at pH 4.0 in detail. F) 150 μM Am-GLP-1 at pH 3.0 and 4.0 in detail.

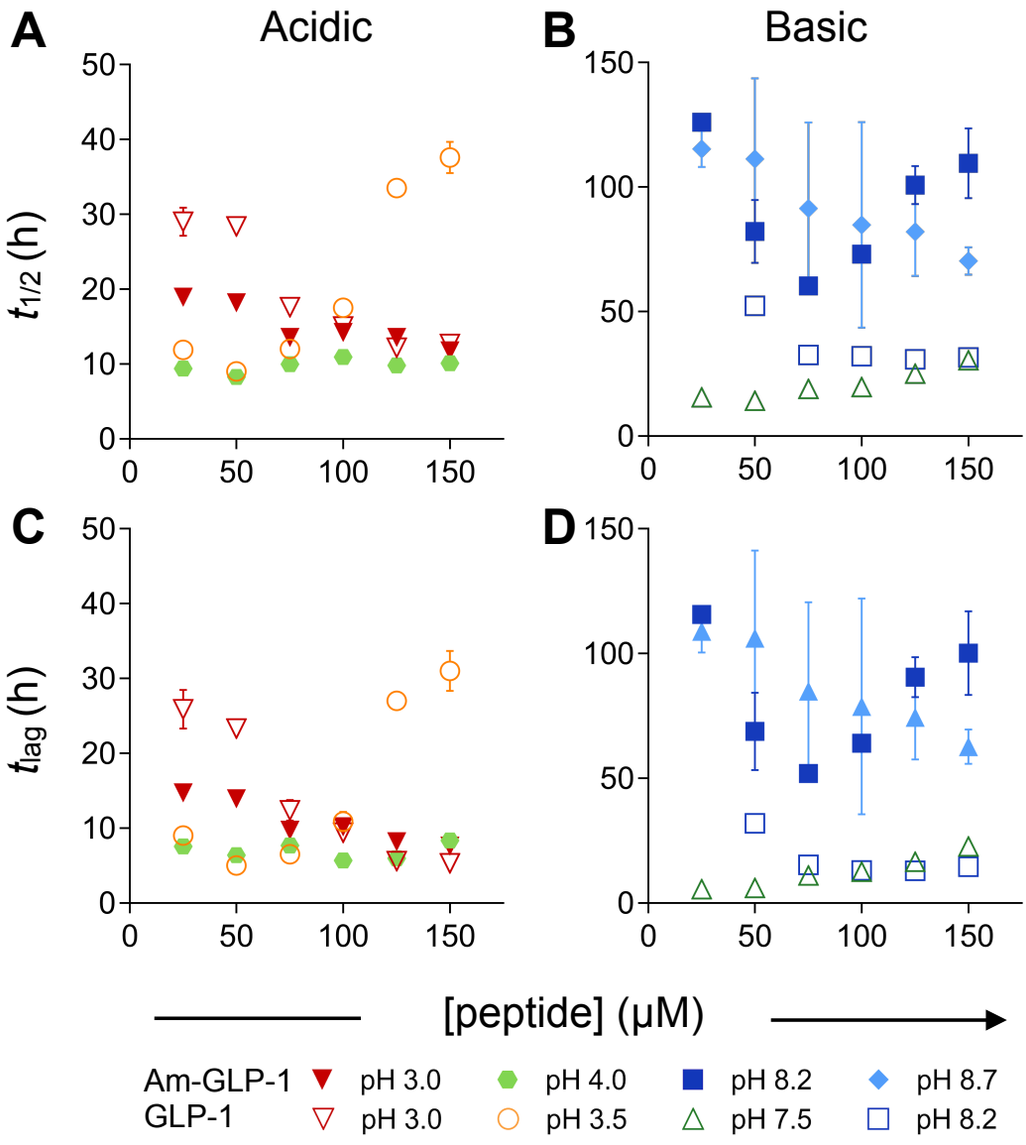


Figure 4.6. Kinetic parameters from the aggregation kinetics of GLP-1 and Am-GLP-1.

The parameters were determined at six different peptide concentrations between 25 – 150 μM. The error bars show the standard deviation. A) and C) $t_{1/2}$ and t_{lag} of GLP-1 at pH 3.0 (red open triangles) and pH 3.5 (orange open circles) as well as Am-GLP-1 at pH 3.0 (red closed triangles) and pH 4.0 (green closed hexagons). All samples in 25 mM citrate buffer. B) and D) $t_{1/2}$ and t_{lag} of GLP-1 at pH 7.5 in phosphate buffer (green open triangles) and Tris at pH 8.2 (dark blue open squares) as well as Am-GLP-1 at pH 8.2 (dark blue closed squares) and pH 8.7 (blue closed diamonds). GLP-1 data measured by Dr Karolina Zapadka.[5]

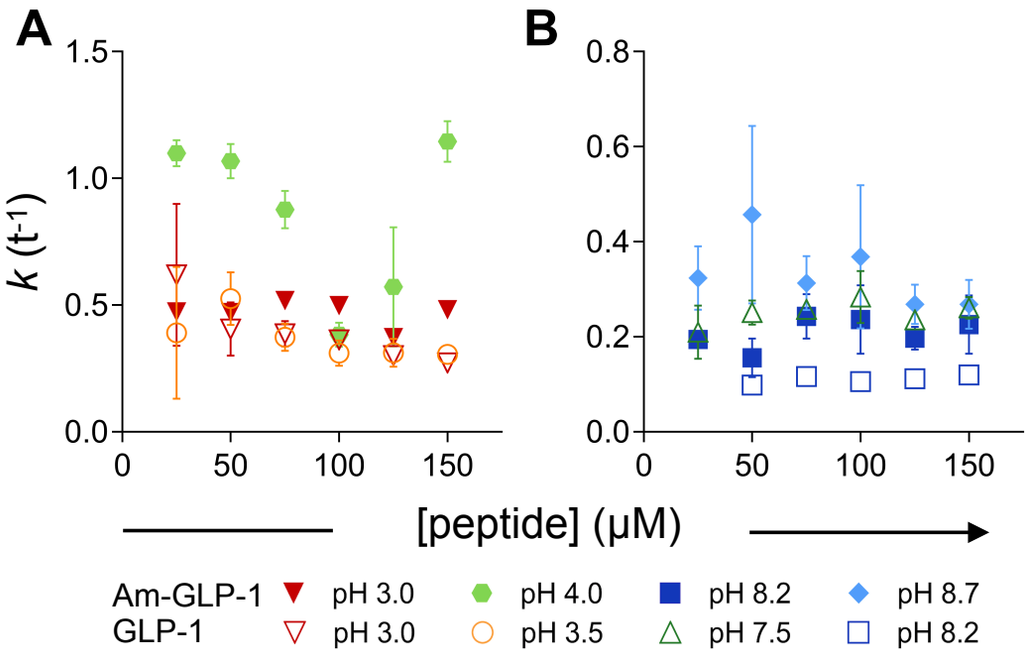


Figure 4.7. Apparent growth rate results.

A) Apparent growth rate (*k*) of GLP-1 at pH 3.0 (red open triangles) and pH 3.5 (orange open circles) as well as Am-GLP-1 at pH 3.0 (red closed triangles) and pH 4.0 (green closed hexagons). All samples in 25 mM citrate buffer. B) Apparent growth rate (*k*) of GLP-1 at pH 7.5 in phosphate buffer (green open triangles) and Tris at pH 8.2 (dark blue open squares) as well as Am-GLP-1 at pH 8.2 (dark blue closed squares) and pH 8.7 (blue closed diamonds). Error bars show the standard deviation between triplicate runs for Am-GLP-1. The apparent growth rates are taken from the sigmoidal fit of the aggregation kinetics shown in Figure 4.5A-D. For GLP-1, the apparent growth rates were measured by Dr Karolina Zapadka.[5]

The data for Am-GLP-1 at pH 8.2 and 8.7 are shown in Figure 4.5C and D, respectively. At pH 8.2, the kinetic data are typical of a nucleation-polymerisation mechanism where lag time decreases with increasing peptide concentration up to a concentration of 100 μM. Above this concentration an increase in $t_{1/2}$ and lag time with increasing peptide concentration is observed, Figure 4.6C & D. The shape of the ThT fluorescence data can be described by a sigmoidal curve with a slight linear increase in fluorescence in the final plateau region. The lag phase itself is completely flat and does not show any increase in fluorescence before the growth phase, Figure 4.5C. At the most basic pH measured, pH 8.7, the variation of the data between runs is pronounced, Figure 4.6C & D.

The variation of the apparent growth rate with peptide concentration is generally low. For both, GLP-1 and Am-GLP-1, the range of values varies by

less than threefold and there is no real trend with peptide concentration, Figure 4.7. Some higher variations in the apparent growth rates are seen for Am-GLP-1 at pH 4.0 and 8.7, but, under these conditions, there are larger standard deviations in these values due to pronounced variations between runs, Figure 4.7A & B.

	pH 3.0		pH 4.0		pH 8.2		pH 8.7	
Conc. (μM)	$t_{1/2}$ (h)	SD	$t_{1/2}$ (h)	SD	$t_{1/2}$ (h)	SD	$t_{1/2}$ (h)	SD
25	19.0	0.0	9.4	0.2	126.0	0.0	115.3	7.2
50	18.2	0.5	8.3	0.3	82.2	12.6	111.3	32.4
75	13.6	0.4	10.0	0.3	60.3	2.2	91.5	34.6
100	14.2	0.2	11.0	0.3	73.0	1.7	84.8	41.2
125	13.6	0.1	9.8	0.6	100.9	7.7	82.0	17.8
150	11.8	0.5	10.1	0.1	109.6	14.0	70.3	5.5
Conc. (μM)	t_{lag} (h)	**SD**	t_{lag} (h)	**SD**	t_{lag} (h)	**SD**	t_{lag} (h)	**SD**
25	14.7	0.1	7.6	0.2	115.7	0.0	108.9	8.5
50	13.9	0.6	6.4	0.3	68.8	15.5	106.2	35.1
75	9.7	0.5	7.7	0.2	51.9	3.5	85.0	35.7
100	10.2	0.2	5.7	0.5	64.1	1.8	78.8	43.2
125	8.2	0.3	6.0	0.9	90.6	8.1	74.4	16.9
150	7.6	0.3	8.3	0.1	100.2	16.8	62.7	6.9
Conc. (μM)	k (h^{-1})	**SD**	k (h^{-1})	**SD**	k (h^{-1})	**SD**	k (h^{-1})	**SD**
25	0.47	0.01	1.10	0.05	0.19	0.00	0.32	0.07
50	0.47	0.02	1.07	0.07	0.16	0.04	0.46	0.19
75	0.52	0.01	0.88	0.07	0.24	0.05	0.31	0.06
100	0.50	0.01	0.38	0.05	0.24	0.07	0.37	0.15
125	0.37	0.01	0.57	0.24	0.20	0.02	0.27	0.04
150	0.48	0.02	1.15	0.08	0.23	0.06	0.27	0.05

Table 4.1. Kinetic parameters from the sigmoidal fits of the Am-GLP-1 ThT assay data shown in Figures 4.5 and 4.6.

SD is the standard deviation.

4.6 The Variance Between Triplicate Runs Increases with Lag Time

At pH 3.0, the observed variance between triplicate runs for Am-GLP-1 is low, however, at higher pH values it increases such that by pH 8.2 the variation is significant and at pH 8.7 it is pronounced, Figures 4.5 and 4.6. The possibility that the increase in variance is caused by impurities in the peptide preparations can be excluded because the same batch of peptide is used for all pH values. The large variance is associated with longer lag times,[352] conditions which favour the population of off-pathway oligomers.

4.7 Additional Analysis of Kinetic Aggregation Data Using a Nucleated-Polymerisation Model

An additional analysis of the ThT assay data of Am-GLP-1 and previously published data of GLP-1[5] was performed by normalising the data and fitting it to an analytical approximation of a nucleated-polymerisation and elongation model with fragmentation (NEF) as a secondary nucleation mechanism, Equation 2.6.[83,353,354] It is assumed that there is no fibril seeding and the complete peptide concentration is monomeric at the beginning of the reaction.

In Equation 2.6, δ and κ represent the effective rate for nucleated polymerisation without secondary pathways (primary nucleation) and the rate of multiplication of the filament population (secondary nucleation), respectively, Section 2.9.3. Both δ and κ are not individually linked to one specific physical process or step in the reaction mechanism but are instead the product of several different steps. The primary nucleation rate, δ, is related to the size of the nucleus, the primary nucleation and elongation rates, as well as dependent upon the total peptide concentration, Equation 2.7. In comparison, the secondary nucleation rate, κ, is related to the primary nucleation rate and the fragmentation rate constant, in addition to being dependent upon the total peptide concentration, Equation 2.8. These relationships are the direct consequence of the solution of a Master equation based on physical principles.[83,353,354]

For Am-GLP-1 and GLP-1, at all four pH values studied, it was not possible to fit the data globally using datasets acquired at different peptide concentrations using constant primary and secondary nucleation rates. Therefore, the data for each peptide concentration were fitted separately to Equation 2.6, Figures 4.8A-D and 4.9A-D.

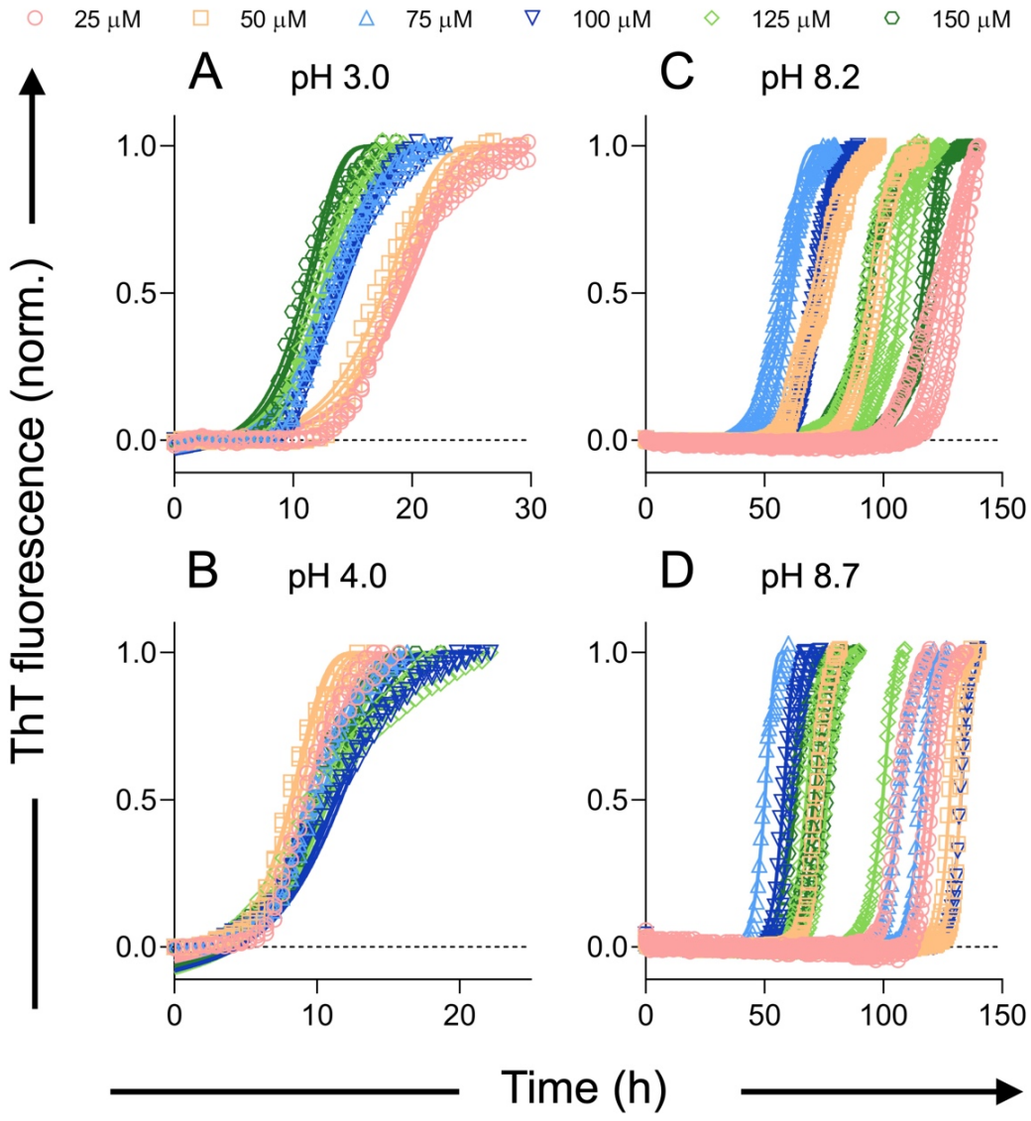


Figure 4.8. Fits of Am-GLP-1 ThT data to the NEF model.

All samples were measured at the following peptide concentrations: 25 μM (pink), 50 μM (orange), 75 μM (blue), 100 μM (dark blue), 125 μM (green), and 150 μM (dark green) and all experiments performed at 37 °C. The data were normalised and fitted using Equation 2.6. The model is based on nucleation, elongation and fragmentation steps. The best fits are shown by the solid lines. All fits showed high R^2 values with the vast majority showing values above 0.97. A) Am-GLP-1 in 25 mM citrate buffer at pH 3.0 B) Am-GLP-1 in 25 mM citrate buffer at pH 4.0. C) Am-GLP-1 in 25 mM Tris buffer at pH 8.2. D) Am-GLP-1 in 25 mM Tris buffer at pH 8.7.

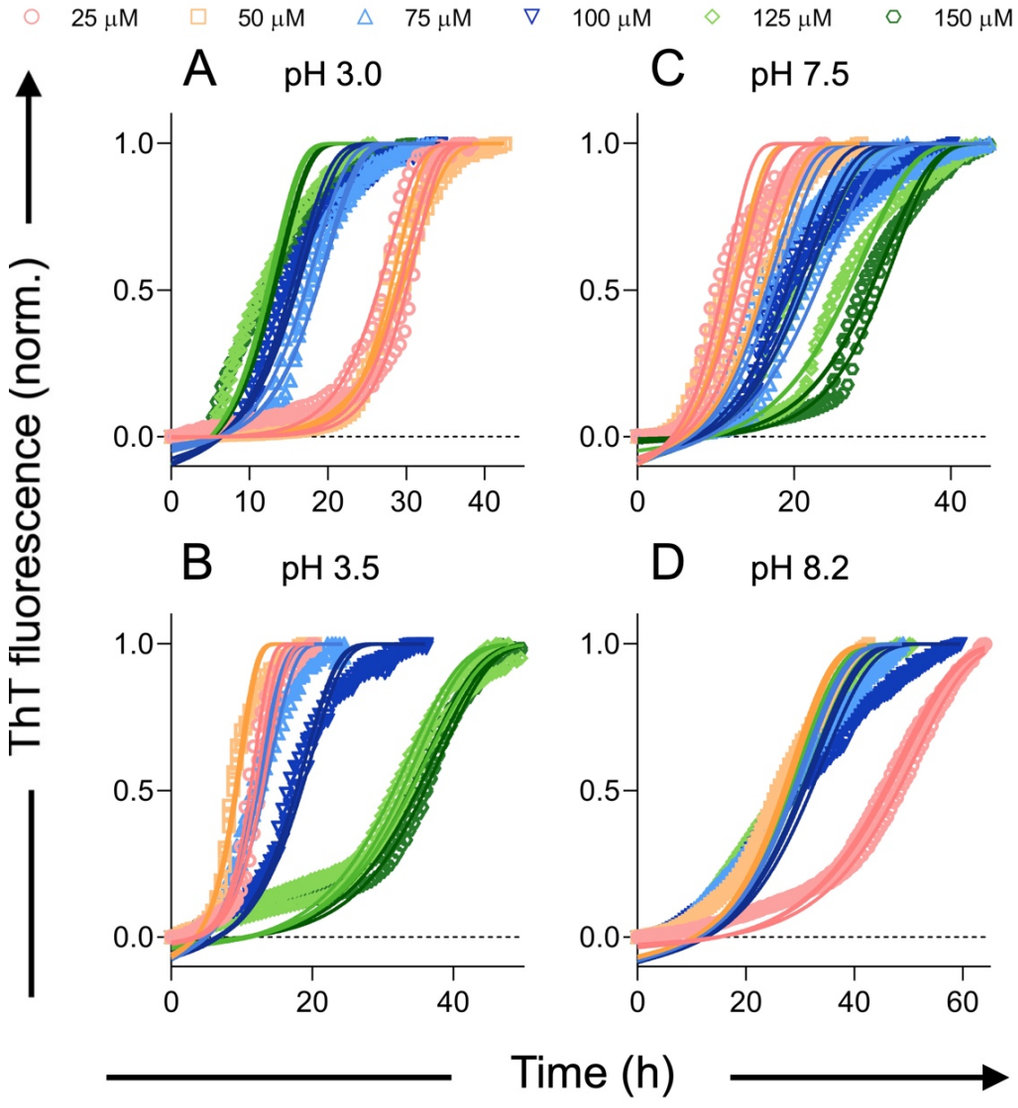


Figure 4.9. Fits of GLP-1 ThT data to the NEF model.

The original data were acquired by Dr Karolina Zapadka. All samples were measured at the following peptide concentrations: 25 μM (pink), 50 μM (orange), 75 μM (blue), 100 μM (dark blue), 125 μM (green), and 150 μM (dark green) and all experiments performed at 37 °C. The data were normalised and fitted using Equation 2.6. The model is based on nucleation, elongation and fragmentation steps. The best fits are shown by the solid lines. All fits showed high R^2 values with the vast majority showing values above 0.97. A) Am-GLP-1 in 25 mM citrate buffer at pH 3.0 B) Am-GLP-1 in 25 mM citrate buffer at pH 4.0. C) Am-GLP-1 in 25 mM Tris buffer at pH 8.2. D) Am-GLP-1 in 25 mM Tris buffer at pH 8.7.

The primary nucleation rates, δ, of Am-GLP-1 are shown in Figure 4.10A & B. δ varies tenfold across pH values for both GLP-1 and Am-GLP-1, but they vary differently with pH. For Am-GLP-1, there is a clear difference in the rates between the samples at acidic and basic pH values. The longer lag times observed under basic conditions are the result of small δ values at basic pH values in comparison to the acidic samples. Consistent with the decrease in the lag time at pH 3.0 with increasing peptide concentration, the primary nucleation rate δ increases with increasing peptide concentration. In contrast, at pH 4.0 the values are approximately constant and independent of peptide concentration.

The values of the secondary nucleation rate, κ, calculated for Am-GLP-1 at all pH values are similar, Figures 4.10C and 4.11B. The values are closely related to the steepness of the apparent growth phase in the ThT data. Although some trends can be observed between κ and peptide concentration, for example, κ increases with increasing peptide concentration at pH 3.0, in general the errors associated with κ are large and it can be considered to be generally independent of peptide concentration. At pH 8.2, the secondary nucleation rate is lower than it is for the acidic samples.

For Am-GLP-1, particularly large variations in δ and κ were calculated at pH 8.7, Figures 4.10C and 4.11A. This may be due to the variance in the maximum ThT signal, Figure 4.5D, which the method assumes is equivalent to the amount of peptide in a fibrillar form capable of binding ThT. The primary nucleation rate even showed negative values at a peptide concentration between 25 and 50 μM which is theoretically not possible but caused by the slight negative slope of the initial baseline in the ThT data, Figure 4.8D. The negative values of the primary nucleation rate are the reason why the data points at 25 and 50 μM disappear in a log-linear plot, Figure 4.11A.

As with the Am-GLP-1 data, the data for GLP-1[5,351] at each pH value and peptide concentration were fitted individually to Equation 2.6. Figure 4.9 shows the best fit of the ThT data to Equation 2.6, whilst the parameters δ and κ from the analysis are shown in Figure 4.10B and D, respectively. The range of values of δ for GLP-1 are similar to Am-GLP-1 but the pH dependence differs. However, some similarities with the Am-GLP-1 data were observed: δ

increases with peptide concentration at pH 3.0. At other pH values, δ is relatively independent of peptide concentration.

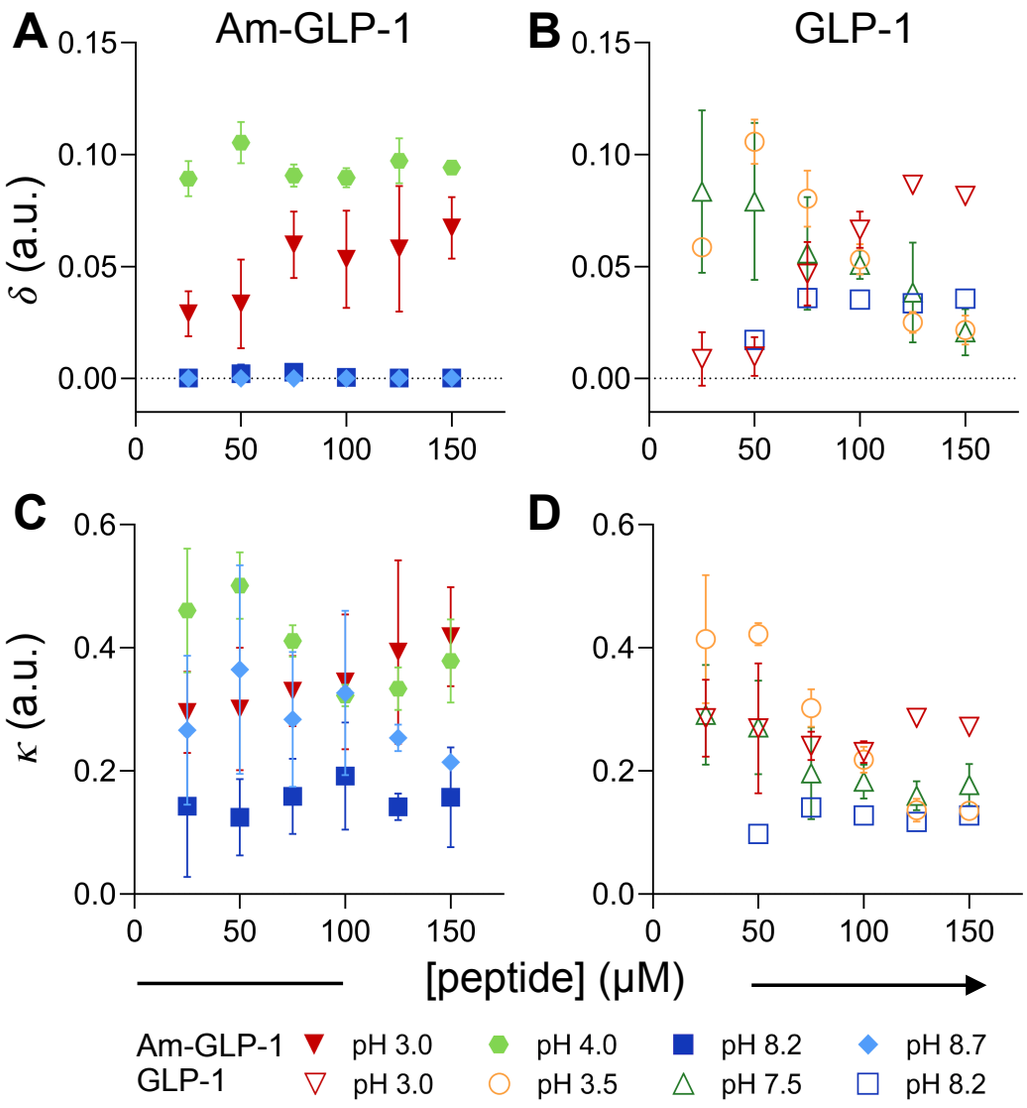


Figure 4.10. Kinetic parameters from the aggregation kinetics of the NEF model of GLP-1 and Am-GLP-1.

The parameters were determined at six different peptide concentrations between 25 – 150 μM. The error bars show the standard deviation of a triplicate measurement. A) Linear plot of the primary nucleation rates, the effective rate for nucleated polymerisation without secondary pathways of Am-GLP-1 at pH 3.0 (red closed triangles), 4.0 (green closed hexagons), 8.2 (dark blue closed squares) and 8.7 (blue closed diamonds). B) Linear plot of the primary nucleation rates, the effective rate for nucleated polymerisation without secondary pathways of GLP-1 at pH 3.0 (red open triangles), 3.5 (orange open circles), 7.5 (green open triangles) and 8.2 (dark blue open squares). C) Linear plot of the secondary nucleation rates i.e., the rate of multiplication of the filament population of Am-GLP-1. D) Linear plot of the secondary nucleation rates i.e., the rate of multiplication of the filament population of GLP-1. GLP-1 data reanalysed from Zapadka *et al.*[5]

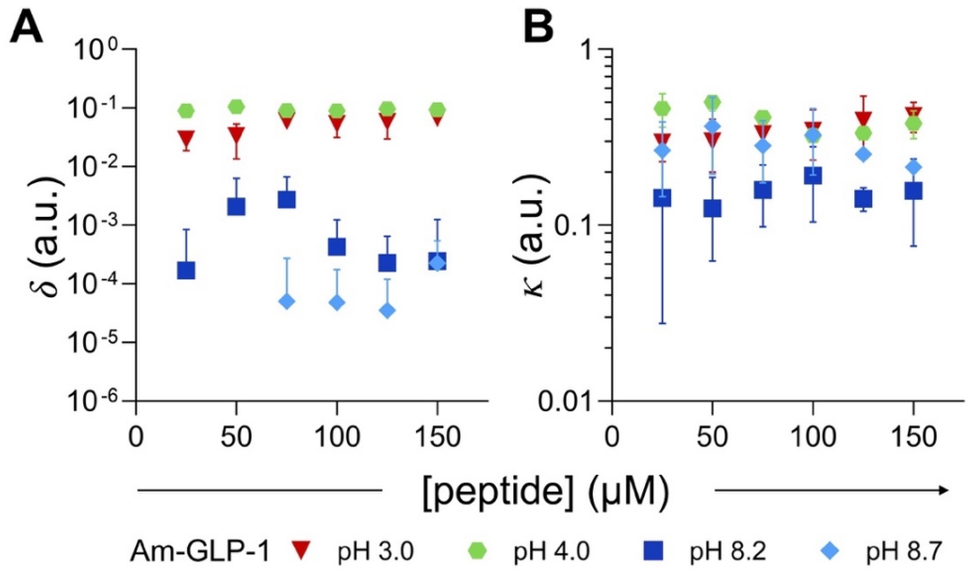


Figure 4.11. Primary and secondary nucleation rates for Am-GLP-1 ThT data fitted to the NEF model.

A) Log-linear plot of the primary nucleation rates, the effective rate for nucleated polymerisation without secondary pathways. B) Linear-log plot of the secondary nucleation rates i.e., the rate of multiplication of the filament population. The data is shown for pH 3.0 (red closed triangles), pH 4.0 (green closed hexagons), pH 8.2 (dark blue closed squares) and pH 8.7 (blue closed diamonds). The error bars show the standard deviation of a triplicate measurement.

The pH dependence of κ is different for GLP-1 and Am-GLP-1. However, the values are in the same range. The data for GLP-1 at pH 3.5 and 7.5 does appear to show a decrease of κ with increasing peptide concentration suggesting that off-pathway events are happening, Figure 4.12. Here too, the error bars are large, but the trend appears to be real.

The analysis of the ThT-assay data of GLP-1 and Am-GLP-1 shows that, if the aggregation kinetics are following a NEF-model, the primary (δ) and secondary (κ) nucleation rates vary with pH as expected. Similar it shows that some of the aggregation parameters like the size of the nucleus, primary nucleation, elongation and fragmentation rates are under most conditions peptide concentration-independent because, under specific conditions, constant δ and κ values have been observed with increasing peptide concentration.[17,274] However, under some conditions, a decrease of κ was detected with increasing peptide concentration.

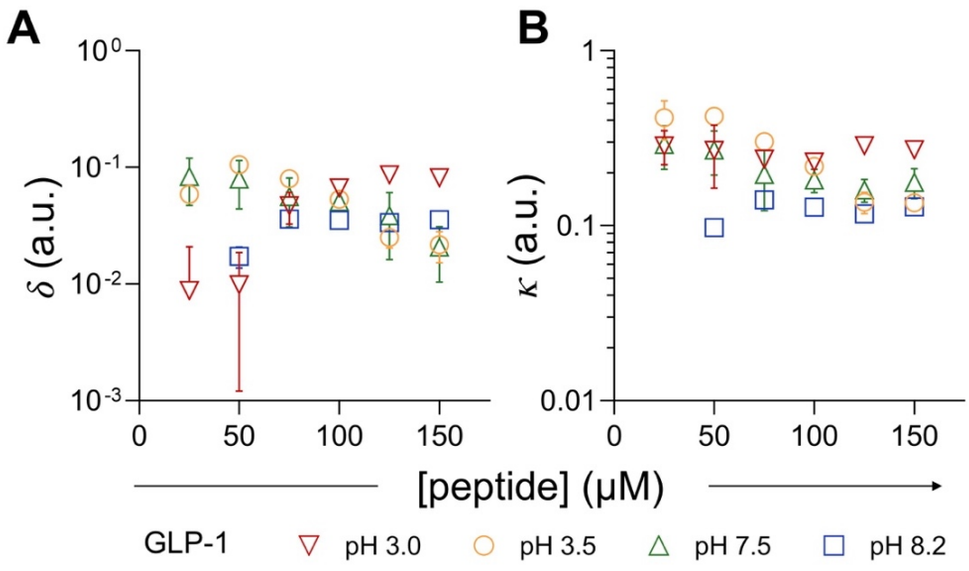


Figure 4.12. Primary and secondary nucleation rates for GLP-1 ThT data fitted to the NEF model.

A) Log-linear plot of the primary nucleation rates, the effective rate for nucleated polymerisation without secondary pathways. B) Linear-log plot of the secondary nucleation rates i.e., the rate of multiplication of the filament population. The data are shown for pH 3.0 (red open triangles), pH 4.0 (orange open circles), pH 8.2 (green open triangles) and pH 8.7 (dark blue open squares). GLP-1 data reanalysed from Zapadka *et al*.[5] The error bars show the standard deviation of a triplicate measurement.

4.8 General Solution to Predict the Maximum Influence of Off-Pathway Species on Monomer Concentration

It is well-established that the kinetics of aggregation of a system that follows a nucleation-propagation mechanism is monomer concentration dependent with higher peptide concentrations leading to a decrease in the lag time. That is why the measured peptide concentration dependence for Am-GLP-1 at pH 4.0, pH 8.2 and pH 8.7 (37 °C), where the lag time is constant or even increases with peptide concentration, are so unusual. It has been suggested that this behaviour is due to the population of off-pathway oligomers.[5] Here, a mathematical proof that the population of off-pathway oligomers cannot lead to an increase in the lag time with increasing peptide concentration is presented.

Assuming that the presence of off-pathway oligomers populated during the aggregation reaction affects the monomer concentration [M], the degree to which the off-pathway species reduce the effective monomer concentration can be estimated. It is possible to show that for two different monomer starting concentrations $[M_1]_0$ and $[M_2]_0$, where $[M_2]_0 > [M_1]_0$, it is true that at every time point the concentration $[M_2](t) > [M_1](t)$. This means that the concentration of monomers accessible to the on-pathway process is at any time always higher in case of $[M_2]$ than $[M_1]$. Mathematically, this becomes clear if the condition $[M_2]_0 > [M_1]_0$ is inserted into the integrated rate law (Equation 4.2), which leads to the Inequation 4.3.

Using an nth order rate law

$$-\frac{d[M]}{dt} = [M]^n \tag{4.1}$$

Integrated rate law

$$\frac{1}{[M]^{n-1}} = \frac{1}{[M]_0{}^{n-1}} + kt \tag{4.2}$$

$$\frac{1}{[M_2]_0{}^{n-1}} + kt < \frac{1}{[M_1]_0{}^{n-1}} + kt \tag{4.3}$$

Following Inequation 4.3, $[M_2](t)$ is always bigger than $[M_1](t)$ if the reaction follows the same rate equation and $[M_2]_0 > [M_1]_0$. It is also independent of the order of the reaction. However, when off-pathway species exist then the amount of monomer available is reduced compared to the amount of monomer that would be populated in the absence of off-pathway oligomers.

4.9 Fibril Morphology and Higher Structures

Because ThT assays provide indirect evidence of fibrillation, scanning electron microscopy (SEM) was used to verify that amyloid-like fibrils of Am-GLP-1 were formed, as well as to gain information on their morphology.

4.9.1 SEM Control Measurements of Pure Buffer

To be able to differentiate between artefacts caused by the buffer or the deionised water used for rinsing and peptide aggregates, controls were undertaken where a glass sheet with buffer only on it was subsequently dried and imaged. Exactly the same method was used as for the protein samples: the samples containing buffer were allowed to dry exactly like the peptide samples, they were also rinsed with deionised water and subsequently dried with a gentle stream of nitrogen.

For all three buffer samples, white round deposits can be observed at lower magnifications, Figures 4.13A, 4.14A and 4.15A. These deposits consist of crystal-like species on the glass sheet as seen for phosphate buffer at pH 3.0, Figure 4.13B and C, or similar crystal-like, sharp but randomly distributed fragments as imaged for phosphate buffer at pH 4.0, Figure 4.14C and D, or Tris buffer at pH 8.2 (37 °C), Figure 4.15D. It is most likely that these fragments are caused by salt residues from the dried buffers. Images without deposits (the uncovered region of the glass microscopy sheet surface) are shown at different magnifications in Figures 4.13D, 4.14B and 4.15B and C. Typically some thin scratches in the glass sheet can be seen.

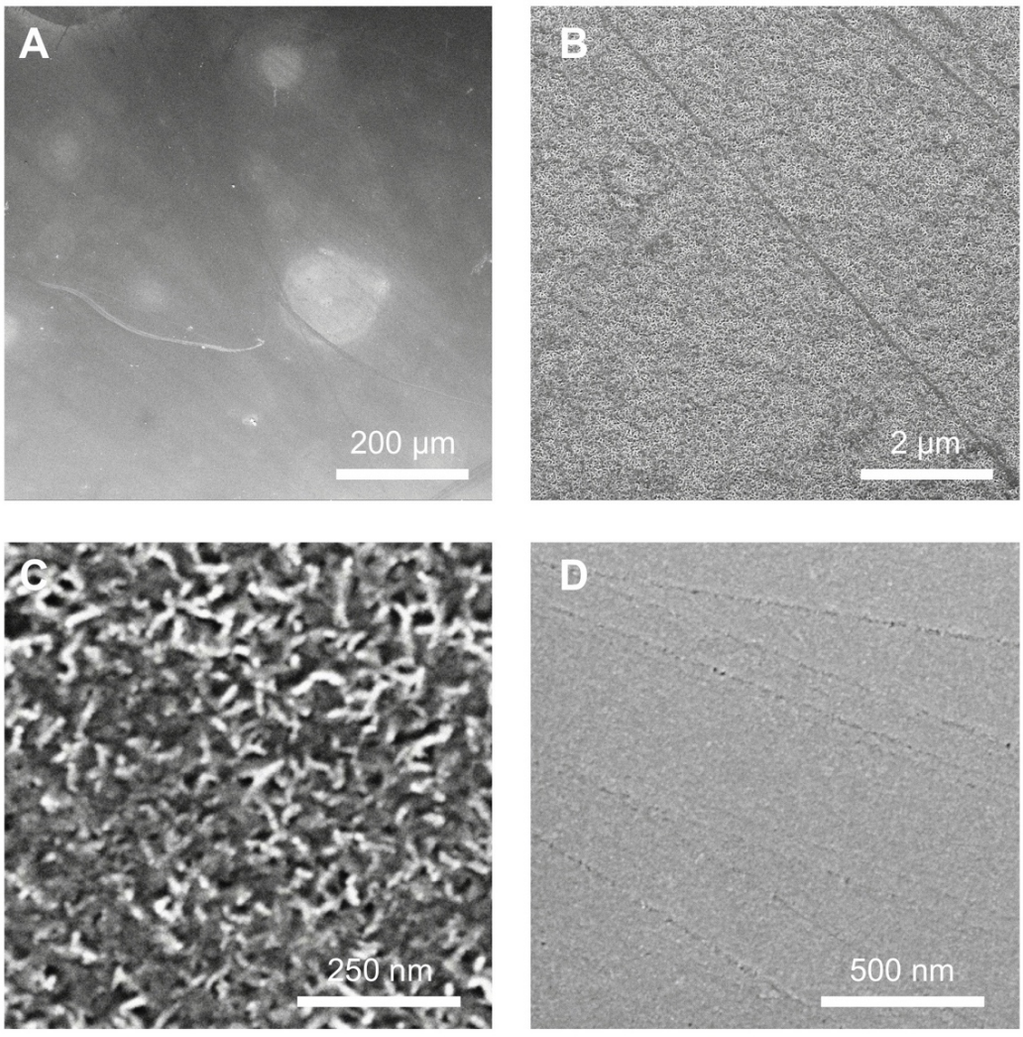


Figure 4.13. SEM images of phosphate buffer at pH 3.0.

The preparation of the samples was exactly the same as for the peptide samples. A) Image at lower magnification shows white deposits and uncovered areas. B and C) Typical images of the deposits seen in Figure 4.13A at different magnifications. D) Images showing the uncovered surface of the glass microscopy slide at a higher magnification.

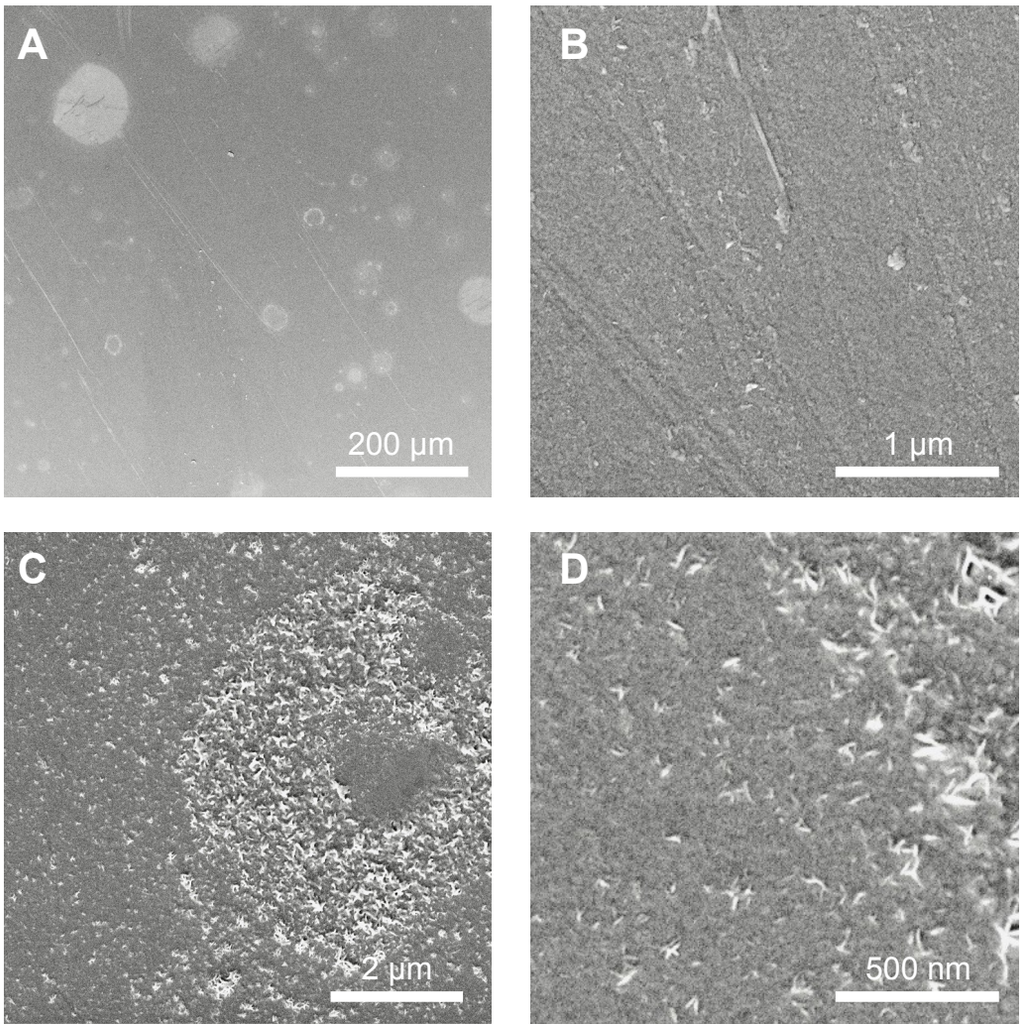


Figure 4.14. SEM images of phosphate buffer at pH 4.0.

A) Image at lower magnification shows the white deposits and uncovered areas. B) Image showing the uncovered surface of the glass microscopy sheet at higher magnification. C and D) Typical images of the deposits seen in Figure 4.14A at different magnifications.

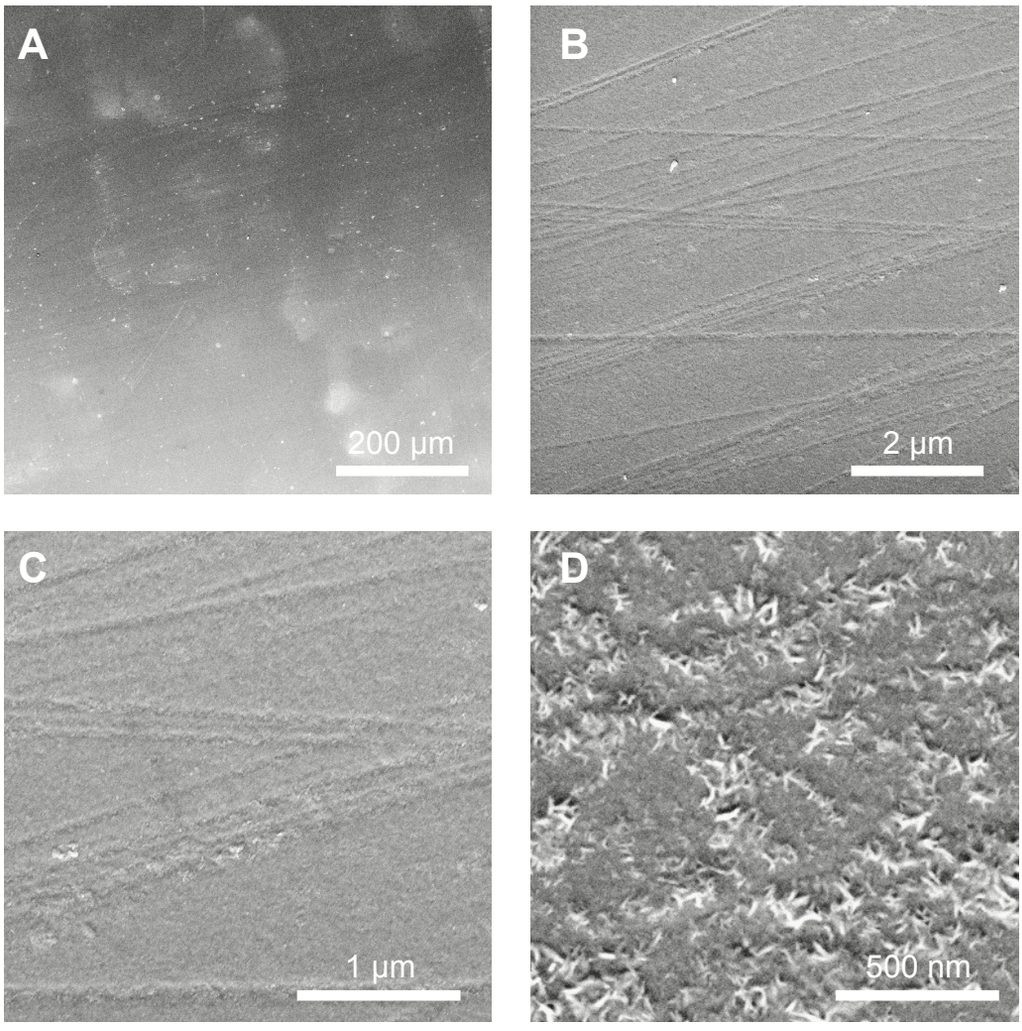


Figure 4.15. SEM images of Tris buffer at pH 8.2.

pH value measured in solution at 37 °C A) Image at lower magnification shows the white deposits and uncovered areas. B and C) Images showing the uncovered surface of the glass microscopy sheet at different magnifications. D) Typical image of the deposits seen in Figure 4.15A at high magnification.

4.9.2 Fibril Morphology and Higher-Order Structure Varies with pH

Samples of Am-GLP-1 at four different pH values, at the end of the ThT aggregation assay, were imaged in three different ways. First, 5 µL of a sample of 150 µM of aggregated Am-GLP-1 was allowed to dry on a microscopy glass sheet. The first spot that was imaged was chosen in a region where there was dense fibril cover. Further images were then taken directly next to the dense fibril cluster where the density of packing was not as high and therefore it was easier to observe single fibrils. In addition, a second 5 µL sample was placed onto an additional glass sheet where it was left for 2 mins before it was gently blown away with nitrogen. This has the advantage that it prevents a high concentration of fibrils covering any other aggregates that might be present. It can, however, favour aggregated species that adhere faster to the glass surface. Typical images of aggregates of Am-GLP-1 formed at pH 3.0, 4.0, 8.2 and 8.7 at 37 °C using the first two methods are shown in Figures 4.16A-D and 4.17A-D, whilst images recorded after the samples were blown dry with nitrogen are shown in Figure 4.18. The SEM images of Am-GLP-1 aggregates were analysed in detail using Fiji (Figures 4.19, 4.20 and 4.21), an image processing package based on ImageJ2. It is worth noticing that SEM imaging does not enable the measurement of the height of the samples. Therefore, the analysis is restricted to distance measurement based on the grey scale of the sample images. Additionally, it is essential to keep in mind that SEM samples have been coated with a 10 nm thick platinum layer.

At pH 3.0, in the most densely covered region, most fibrils are parallel to each other and longer than 2 µm, Figure 4.16A. A few short fibrils are misaligned and spread over the surface in a disordered way. The width of the fibrils is 21 ± 3 nm. In the less densely covered region two different, straight, slightly thicker (width: 42 ± 3 and 61 ± 3 nm), 200-900 nm long fibrils are observed, Figure 4.16B. There are also some shorter fragments with a length of 50-300 nm and 26 ± 4 nm width, which are randomly distributed.

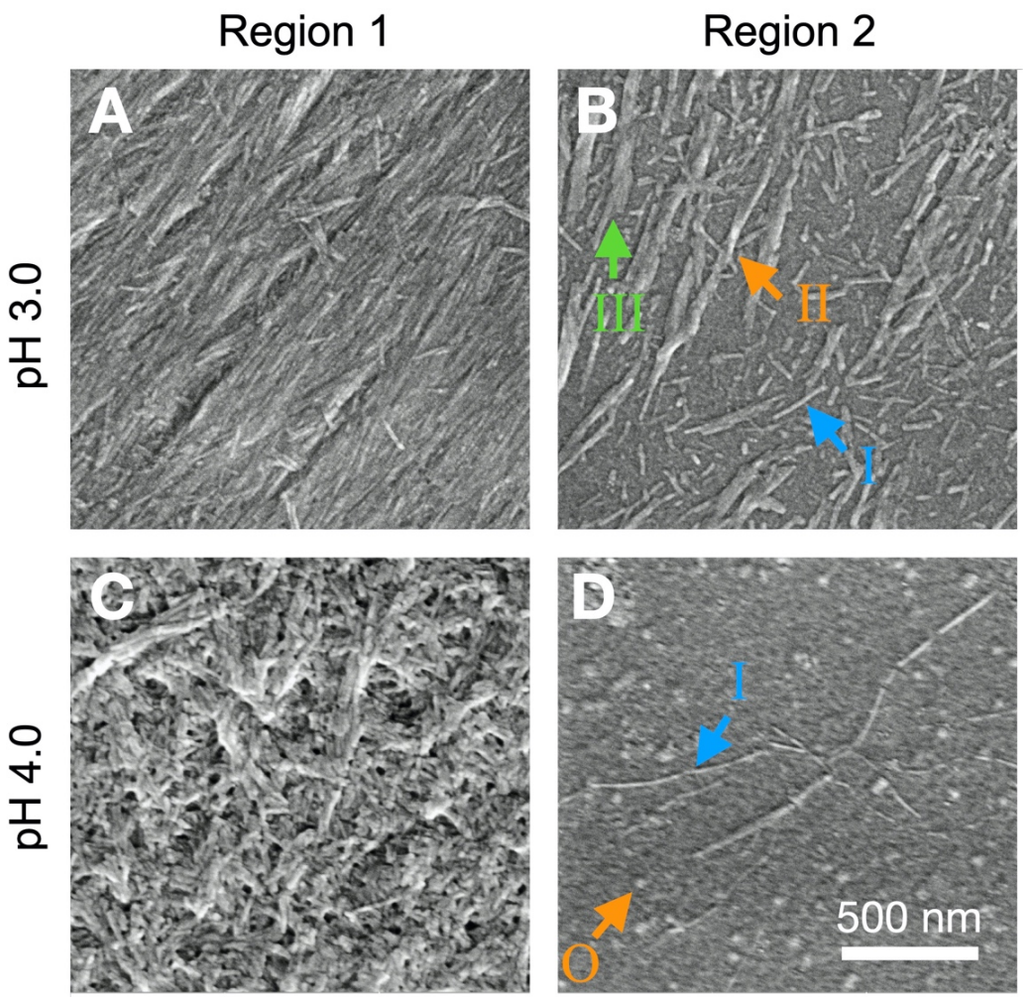


Figure 4.16. SEM images of aggregated forms of Am-GLP-1 under acidic conditions.

After 144 h of incubation time in the plate reader at 37 °C, 5 µL of 150 µM sample were pipetted on to a glass microscopy sheet and subsequently dried. All images have the same magnification. A & B) Am-GLP-1 aggregated in 25 mM citrate buffer at pH 3.0. C & D) Am-GLP-1 aggregated in 25 mM citrate buffer at pH 4.0. The images show the most typical of those obtained. A & C) show images of regions where there is a high local concentration of fibrils (Region 1). B & D) show images of less densely populated regions of the surface (Region 2). B) Three different types of fibrils were observed and analysed. These are indicated by coloured arrows (I blue, diameter approx. 26 nm, II orange, diameter approx. 42 nm, and III, green, diameter approx. 61 nm) D) Typical fibril class I (blue arrow, diameter approx. 22 nm) and oligomeric structure O (orange arrow, diameter between 10 - 50 nm).

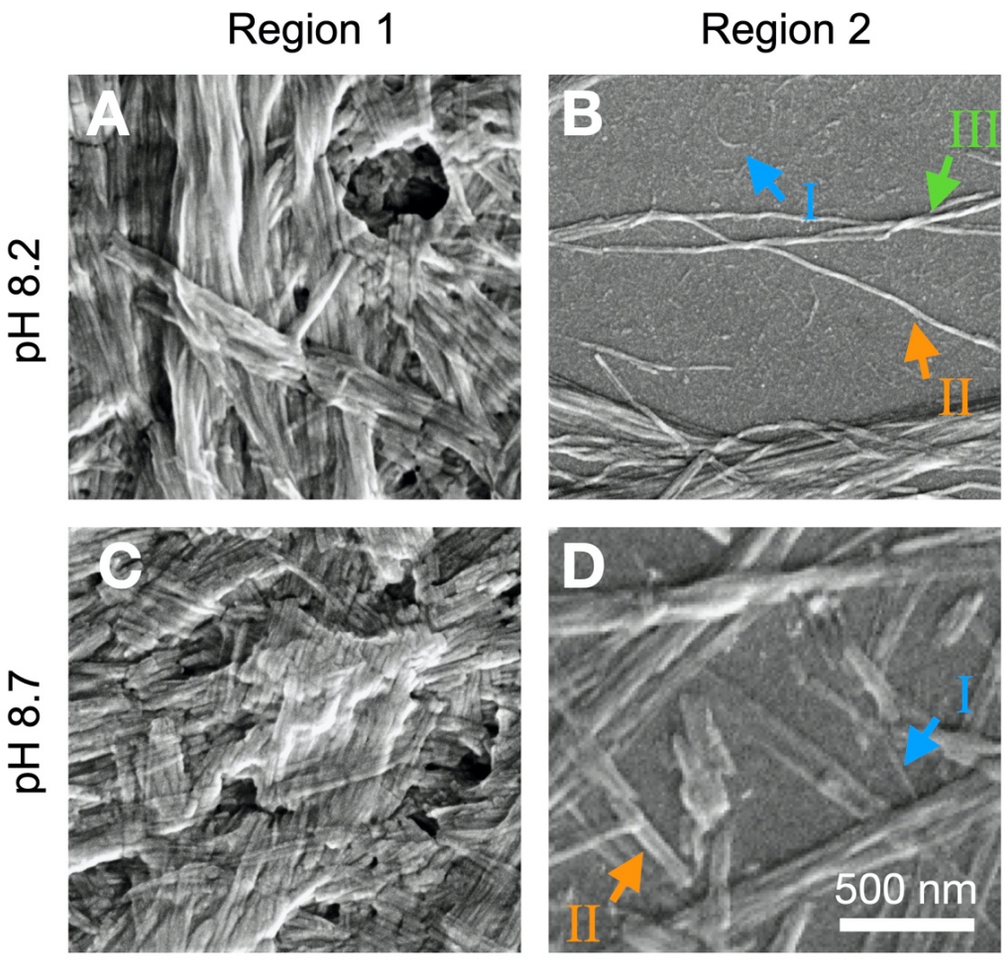


Figure 4.17. SEM images of aggregated forms of Am-GLP-1 under basic conditions.

After 144 h of incubation time in the plate reader at 37 °C, 5 µL of 150 µM sample was pipetted on to a glass microscopy sheet and subsequently dried. All images have the same magnification. A & B) Am-GLP-1 aggregated in Tris buffer at pH 8.2. C & D) Am-GLP-1 aggregated in Tris buffer at pH 8.7. The images show the most typical of those obtained. A & C) show images of regions where there is a high local concentration of fibrils (Region 1). B & D) show images of less densely populated regions of the surface (Region 2). B) Three different types of fibrils are observed and were analysed (I/II/III). These are indicated by coloured arrows (I blue, diameter approx. 20 nm, II orange, diameter approx. 24 nm, and III, green, diameter approx. 30 nm). D) Fibrils type I (blue, diameter approx. 31 nm) and II (orange, diameter approx. 76 nm).

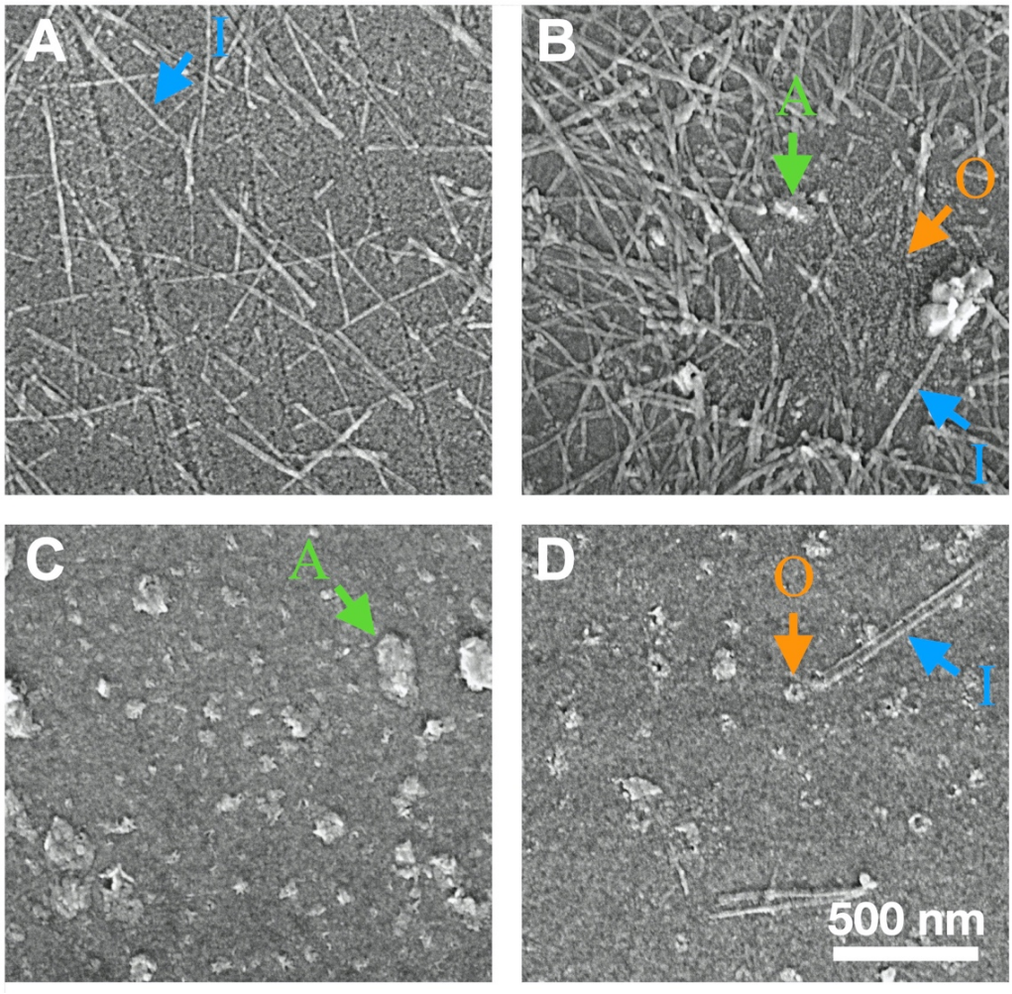


Figure 4.18. SEM images of aggregated forms of Am-GLP-1 samples dried with nitrogen.

After 144 h of incubation time in the plate reader at 37 °C, 5 µL of 150 µM sample were pipetted on to a glass microscopy sheet and the drop was blown away with a gentle nitrogen stream after 2 min on the surface. Different types of analysed aggregates are marked with an arrow. All panels are at the same magnification. Different species are labelled and indicated with arrows. A) Am-GLP-1 aggregated in 25 mM citrate buffer at pH 3.0. Typical fibril type I indicated with a blue arrow (diameter approx. 20 nm). B) Am-GLP-1 aggregated in 25 mM citrate buffer at pH 4.0. Fibrils type I (blue, diameter approx. 22 nm), oligomer-like aggregates are marked O in orange and amorphous aggregates marked A in green. C) Am-GLP-1 aggregated in Tris buffer at pH 8.2. Amorphous aggregates marked A in green. D) Am-GLP-1 aggregated in Tris buffer at pH 8.7. Fibril type I (blue, diameter approx. 53 nm) and circular oligomer O (orange, diameter approx. 70 nm) are marked.

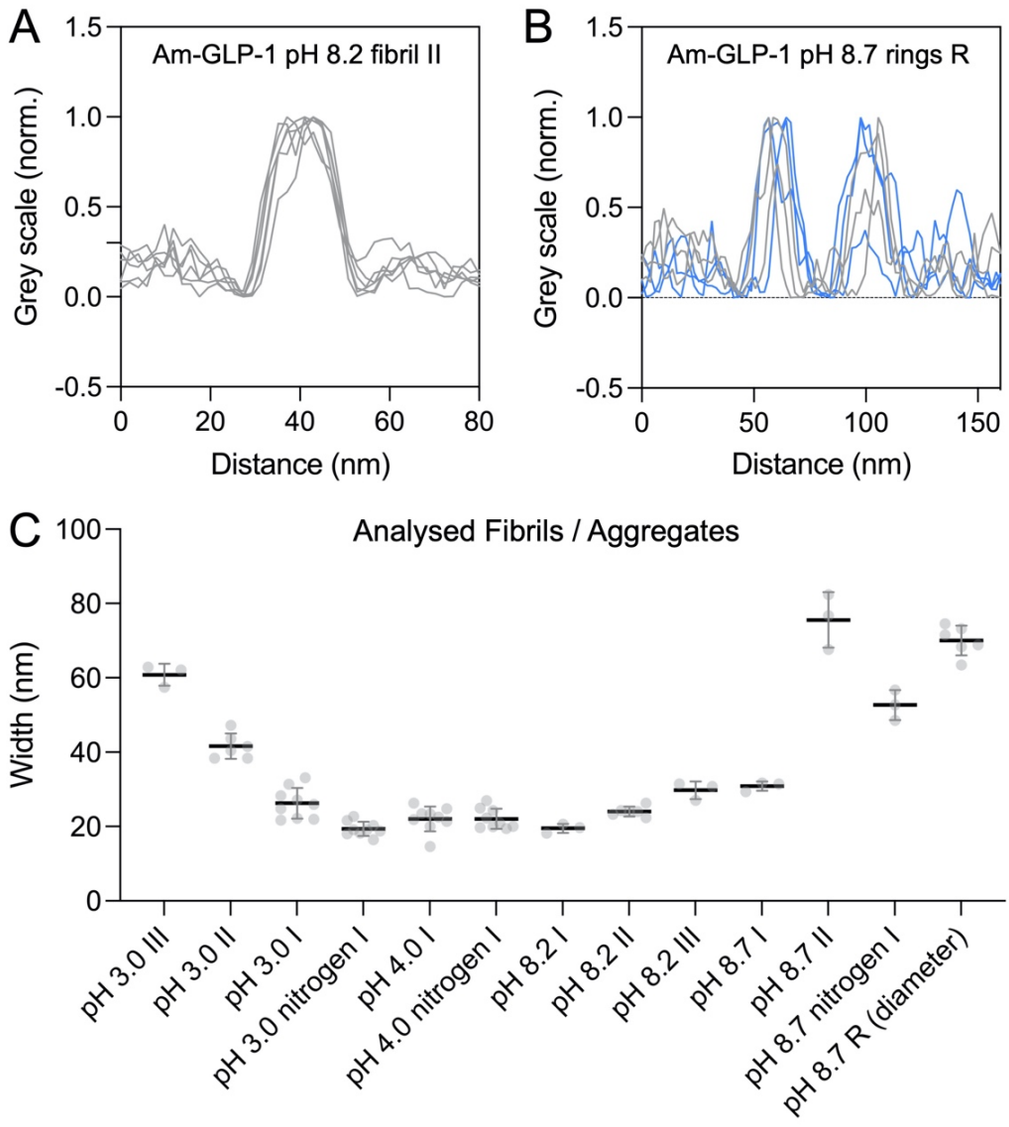


Figure 4.19. Width or diameter of Am-GLP-1 fibrils and oligomers determined using grey scale profiles.

The grey scale profiles were analysed using Fiji. A) A typical example of grey scale profiles to determine the widths of fibrils. Am-GLP-1 incubated at pH 8.2 at 37 °C. B) Grey scale profiles of two circular or ring-like oligomers. For each oligomer three profiles at three different angles were taken. The profiles of Oligomer A and B are shown in blue and grey, respectively. Am-GLP-1 incubated at pH 8.7 at 37 °C. C) Width or diameter of eleven different aggregate types as observed in Figures 4.16, 4.17 and 4.18.

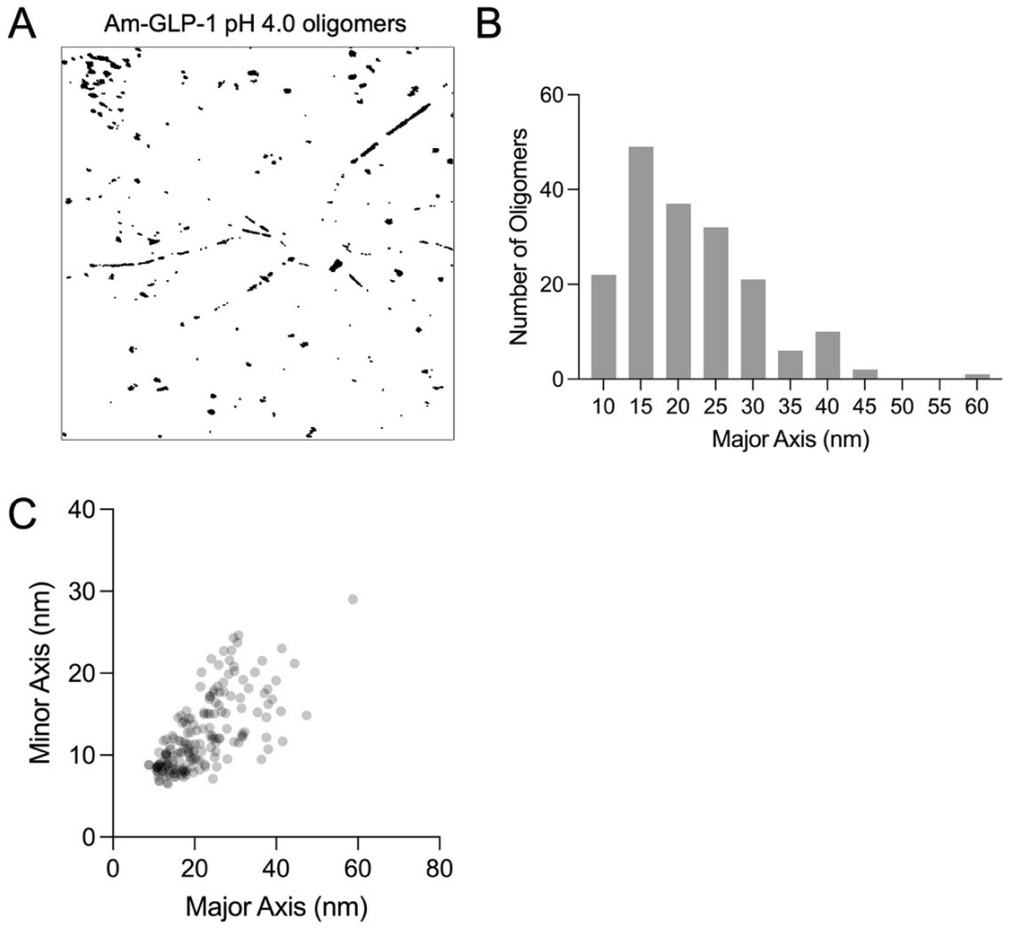


Figure 4.20. Measurement of major and minor axis of Am-GLP-1 oligomers formed in 25 mM citrate buffer at pH 4.0.

A) Binary image of Figure 4.16D. B) Histogram of the oligomer major axis size distribution. To avoid counting fibril fragments, only aggregates with a circularity value higher than 0.5 were considered. The distribution follows a Gaussian with a mean value of 18.5, a standard deviation of 9.1 and R^2 = 0.93. C) Relation between minor and minor axis of the detected oligomers.

For the nitrogen-dried samples, thin fibrillar species (approx. diameter 19 ± 2 nm, within error of that measured for the type I fibrils observed in the earlier experiment), can be seen. These are approx. 200-500 nm in length, Figure 4.18A. Overall, at pH 3.0, it appears as if high concentrations of fibrils of Am-GLP-1 have a tendency to stick to each other, as shown by the fibrils with diameters which are twice and three times the size of the thinnest fibrils observed.

The morphology of Am-GLP-1 fibrils at pH 4.0 shows different patterns. In regions of high coverage, the fibrils are dense but form a randomly distributed mesh appearing sponge-like, with little holes in between, which consists of 15-25 nm thick and 30-200 nm long fibrils, Figure 4.16C. Smaller oligomers with

a diameter of 10 - 45 nm (Figure 4.20A-C) and some longer thin fibrils (22 ± 3 nm wide and 500-1000 nm long) were typically found next to the thick patch of fibrils, Figure 4.16D. From the analysis, it becomes clear that the oligomers are not perfectly round as the minor and major axis lengths stand in a relation of approximately 0.35, Figure 4.20C. For the nitrogen-dried sample, similar thin fibrils were observed as well as oligomer-like species and larger aggregates, Figure 4.18B.

At pH 8.2, the image taken in the region with high fibril coverage shows that the fibrils appear to organise into bundles (Figure 4.17A). In less densely covered regions, the fibrils can be seen more easily and are twisted.

At least two different species have been measured, which are either 24 ± 1 nm or 30 ± 2 nm thick. The total length of one turn of the thinner fibril is slightly shorter than 200 nm, whereas the thicker fibril turns with a periodicity of 85 ± 10 nm, Figure 4.21. Most of these fibrils are longer than 2 µm and, therefore, longer than those formed under acidic conditions, Figure 4.17B.

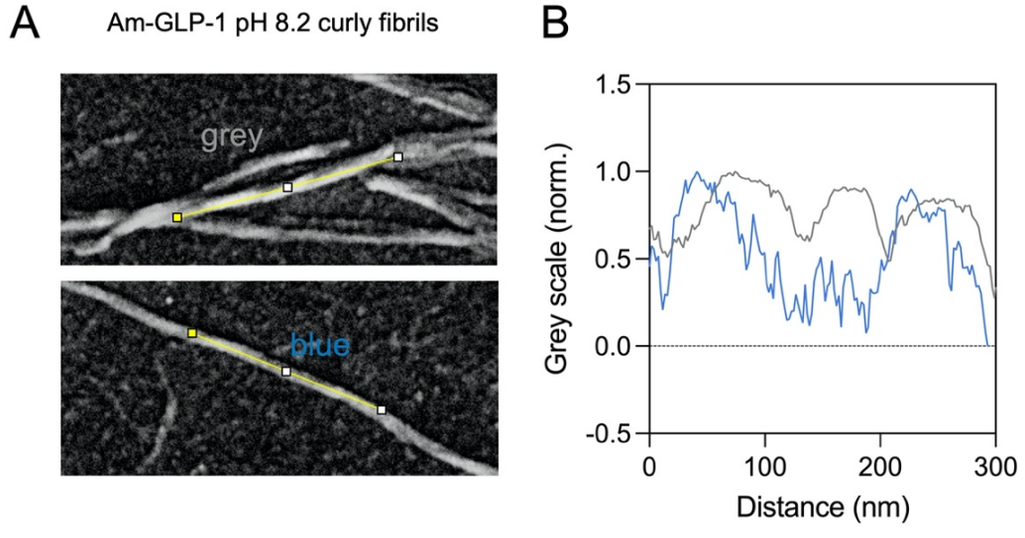


Figure 4.21. Grey scale profiles along Am-GLP-1 fibrils.

A) Am-GLP-1 fibrils formed in 25 mM Tris buffer at pH 8.2 (37°C). B) Grey scale profiles of curly fibrils II (grey) and III (blue) along their longitudinal axis. They show a turn length of 85 ± 10 nm (grey) and slightly less than 200 nm (blue) which has been measured using the distance of the maximum amplitudes.

There are also some thin (width: 20 ± 1 nm), slightly curved fibrils, which are shorter than 100 nm, Figure 4.16B. The image of the sample dried with nitrogen shows a high number of oligomer-like species or really short fibril fragments which are not thicker than 10 nm and only 20-40 nm long. Some of them appear to arrange into larger clusters with a width up to 200 nm, Figure 4.18C.

Fibrils of Am-GLP-1 formed at pH 8.7 appear to be similar in terms of length and width to those obtained at pH 8.2. They appear, however, slightly less rigid and allow covered fibrils to trace through the top layer of fibrils especially if they cross each other as shown in Figure 4.17C. In the less dense region, Figure 4.17D, they also form bundle-like structures with a width of 50-200 nm that are mostly longer than 500 nm. Donut-like shaped oligomers with a diameter of 70 ± 4 nm can be seen in the nitrogen-dried sample shown in Figures 4.18D and 4.19B.

The SEM pictures of Am-GLP-1 show the whole complexity of the aggregation reaction and the different higher-order structures that can be formed at different pHs. Two other main observations are striking. First, the images taken from glass sheets on which the sample drop was blown away after 2 min show differences in the ability of different species to stick to the glass surface. Am-GLP-1 aggregated under acidic conditions shows fibrils sticking to the glass sheet within two minutes, Figure 4.18A & B, whereas both glass sheets covered with sample aggregated under basic conditions, Figure 4.18C & D, did not show fibrils after the two minutes but instead had oligomers covering the surface. This is possibly driven by the different net charge on the peptide and therefore on the aggregated species formed under acidic and basic conditions. Second, only samples with dominant off-pathway species populated during aggregation, as shown by the aggregation kinetics, show significant amounts of oligomeric species, as is the case of Am-GLP-1 at pH 4.0, 8.2 and 8.7 (Figures 4.16D and 4.18C, D).

In our previous study, aggregates of GLP-1 were imaged using atomic force microscopy.[5,351] Similar results were obtained, with the most common fibrils forming structures around 15-25 nm in width or multiples of this. The results of the SEM data for Am-GLP-1 reported here and the previously published AFM data for GLP-1 are summarised in Table 4.2 and Table 3.2, respectively.

pH	Region 1	Region 2	Nitrogen dried
3.0	**Fibrils** width 15-25 nm length: > 2000 nm	**Fibrils a)** width: 26 ± 3 nm length: 50-300 nm fragmented, straight **Fibrils b/c)** width: 42 ± 3 / 61 ± 3 nm length: 200-900 nm straight	**Fibrils** width: 19 ± 2 nm
4.0	**Fibrils** width: 15-25 nm length: 30-200 nm spongy / cavernous / porous	**Oligomers** diameter: 10-45 nm Mean value of major axis: 19 ± 9 nm **Fibrils** width: 22 ± 3 nm length: 500-1000 nm	**Oligomer like dots** diameter: 8-15 nm **Fibrils** width: 22 ± 3 nm length: > 500 nm
8.2	**Fibrils** width: 15-25 nm length: >2000 nm twisted	**Fibrils a)** width: 20 ± 1 nm length: < 100 nm straight **Fibrils b/c)** width: 24 ± 1 / 30 ± 2 nm length: > 2000 nm straight, twisted	**Oligomer like fragments** width: 10 nm thick length: 20-40 nm **Bigger clusters** consisting out of the fragments diameter: < 200 nm
8.7	**Fibrils** width: 15-25 nm length: > 2000 twisted, similar to pH 8.5 but they seem to break more easily in general or during the drying process especially if they cross other fibrils	**Fibril bundles** width: 50 - 200 nm length: > 500 nm slightly twisted	**Oligomers** diameter: 70 ± 4 nm Donut shaped **Fibrils** width: 53 ± 4 nm

Table 4.2. Description of Am-GLP-1 aggregates imaged by SEM

4.10 Conformational Differences Between Aggregates Formed Under Acidic and Basic Conditions

The results from SEM imaging showed that Am-GLP-1 forms different high-order structures at different pH values. For a deeper insight into these structures, ANS and tryptophan fluorescence experiments were conducted for both peptides to probe whether a) hydrophobic patches are exposed and b) the environment around the tryptophan, in the aggregates.

ANS is known to bind to hydrophobic patches in proteins and peptide aggregates and increase its fluorescence.[350] Immediately after the peptide was dissolved and the sample filtered, the ANS fluorescence signal is low, Figure 4.22A. The same samples measured after 144 h of incubation at 37 °C are shown in Figure 4.22B. All basic and neutral samples show no change after aggregation, however, all four acidic samples show a clear increase in fluorescence signal indicative of ANS binding. This suggests that under acidic conditions species are formed which have hydrophobic patches exposed and binding of ANS is favourable. It is possible that the different results obtained at acidic *versus* neutral/basic pH may be due to the difference in the net charge of the peptides under these conditions. ANS, which is negatively charged, binding is likely to be more favourable to peptide aggregates at low pH values where the peptide is positively charged under the conditions. Thus, it is difficult to be quantitative about the relative amounts of aggregated species with exposed hydrophobic patches present at different pH values. Interpretation of these data is also not straightforward as, although it is frequently assumed that ANS binds specifically to oligomers, experiments shown in Section 4.11 show that it also binds to fibrils under certain conditions.

The tryptophan fluorescence data shows that, before the aggregation assay, the signals of all samples are similar with a wavelength of maximum fluorescence (λ_{max}) around 350 nm, Figures 4.22C & D and 4.23. At pH 3.0 and 4.0, the wavelength of maximum fluorescence changes from 350 to 335 nm over the time course of aggregation compared to a much smaller change from 350 to 345 nm at pH 8.2 and 8.7, Figure 4.23. A shorter wavelength for λ_{max} is known

to indicate that the tryptophan is more buried and has less contact with the aqueous solvent.[355]

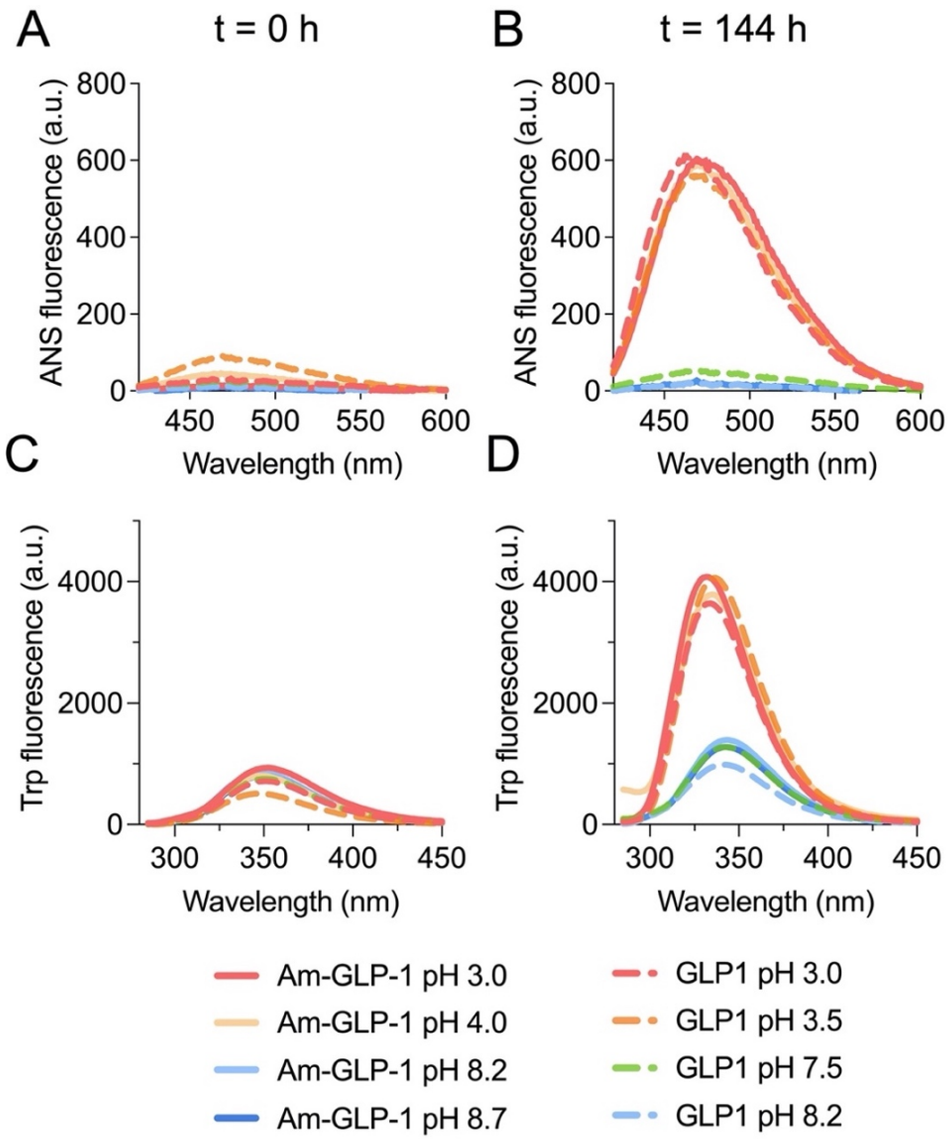


Figure 4.22. ANS and Trp fluorescence of Am-GLP-1 and GLP-1 before and after aggregation

ANS binding and tryptophan (Trp) assay of 150 µM Am-GLP-1 and GLP-1 samples at room temperature. Am-GLP-1 incubated at 37 °C in citrate buffer at pH 3.0, citrate buffer pH 4.0, Tris buffer at pH 8.2 and Tris buffer at pH 8.7. GLP-1 incubated at 37 °C in citrate buffer at pH 3.0, citrate buffer at pH 3.5, phosphate buffer at pH 7.5 and Tris buffer at pH 8.2. The pH of the Tris buffers used during the fluorescence measurements changed from pH 8.2 to 8.5 and pH 8.7 to 9.0 because of the temperature change from 37 °C (aggregation assay) to room temperature (spectroscopic analysis). The buffer concentration was 25 mM. A) ANS fluorescence of Am-GLP-1 and GLP-1 samples before the aggregation assay. B) ANS fluorescence of Am-GLP-1 and GLP-1 samples after aggregation. C) Trp fluorescence of Am-GLP-1 and GLP-1 samples before the aggregation assay. D) Trp fluorescence of Am-GLP-1 and GLP-1 samples after aggregation.

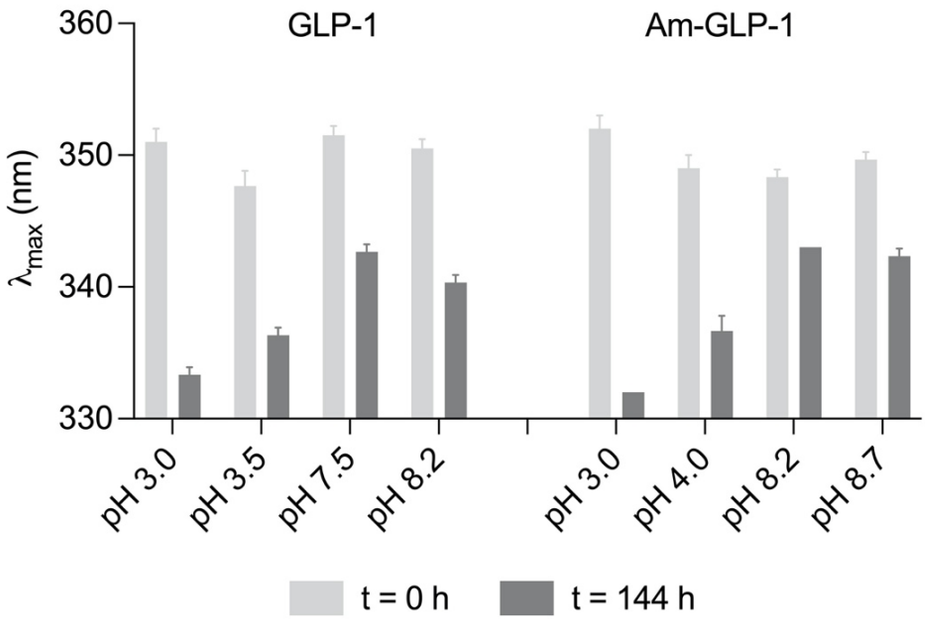


Figure 4.23. Tryptophan fluorescence: shift in λ_{max} after aggregation

λ_{max} of tryptophan fluorescence of Am-GLP-1 and GLP-1 before and after the aggregation assay, Figure 4.22.

4.11 Specificity of ANS Binding to Oligomers/Fibrils

ANS is known to bind to hydrophobic patches and one common assumption is that ANS binds only to oligomeric intermediates and not to monomers or fibrils.[356–358] If this is true, it would enable the measurement of the relative amount of oligomers in different samples at different pH values and times. However, there is some evidence that ANS might also bind, albeit more weakly, to fibrils. Consequently, it is important to establish whether the increase in ANS fluorescence measured at the end of the aggregation assays is caused by oligomers in equilibrium with fibrils or due to direct binding of the ANS to the fibrils. To test this, an aggregation assay (without ThT which can interfere with ANS fluorescence) with GLP-1 and Am-GLP-1 was run. At the end of the incubation period, part of the sample was mixed with ANS or ThT and the fluorescence measured, in addition to the tryptophan fluorescence. The other half of the sample was centrifuged, and the pellet carefully separated from the supernatant, this was washed with buffer and recentrifuged, then the pellet was resuspended in the same volume of the buffer used initially.

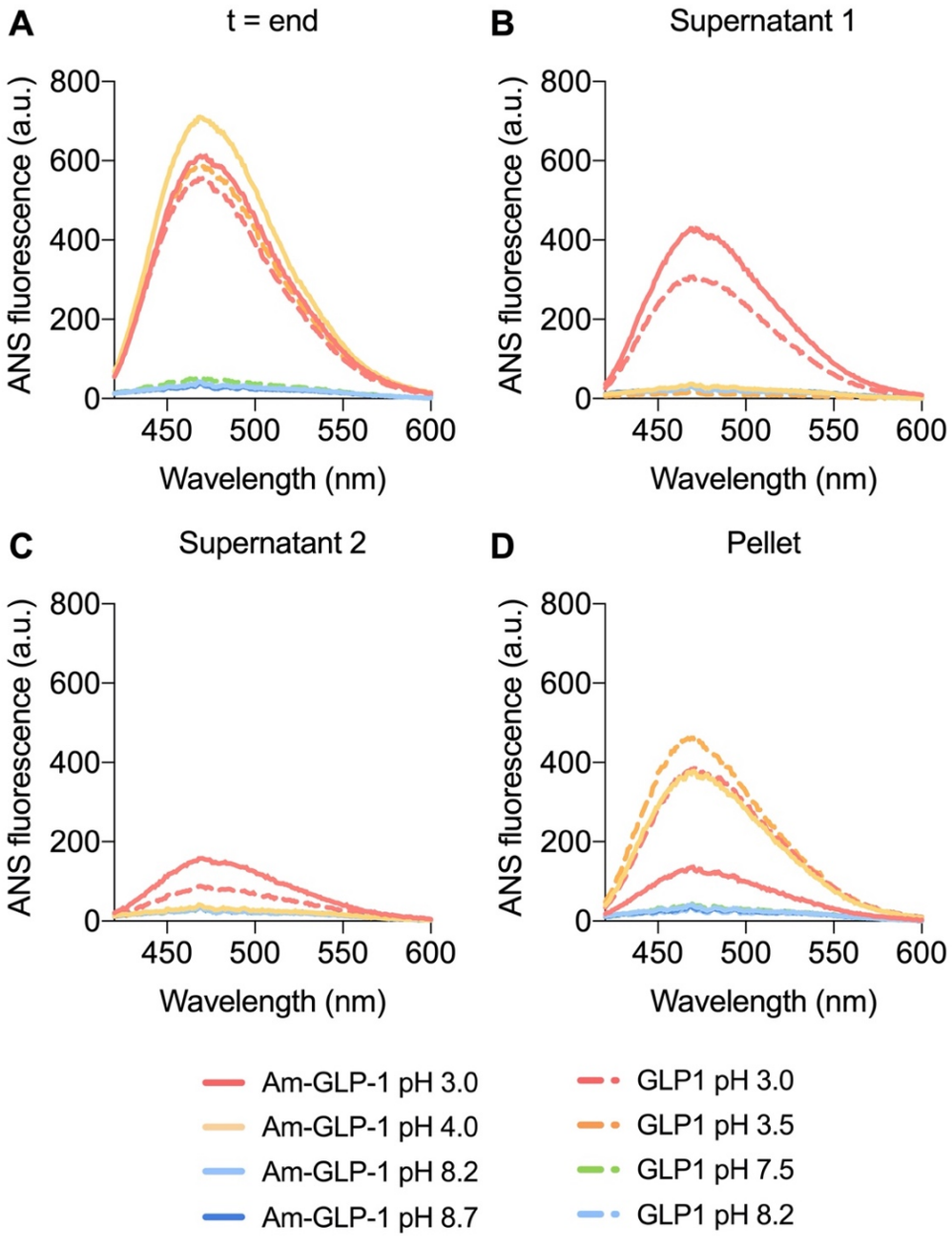


Figure 4.24. ANS fluorescence of GLP-1 and Am-GLP-1 aggregates formed after aggregation and after separation of insoluble/soluble fractions.

ANS binding assay of 150 µM Am-GLP-1 and GLP-1 samples. Am-GLP-1 incubated at 37 °C in citrate buffer at pH 3.0, citrate buffer pH 4.0, Tris buffer at pH 8.2 and Tris buffer at pH 8.7 (9.0 at 25 °C). GLP-1 incubated at 37 °C in citrate buffer at pH 3.0, citrate buffer at pH 3.5, phosphate buffer at pH 7.5 and Tris buffer at pH 8.2 (8.5 at 25 °C). The pH of the Tris buffers used during the fluorescence measurements changed from pH 8.2 to 8.5 and pH 8.7 to 9.0 because of the temperature change from 37 °C (aggregation assay) to room temperature (spectroscopic analysis). All buffers were used at a concentration of 25 mM. A) ANS fluorescence after 140 h of incubation at 37 °C in plate reader B) ANS fluorescence of supernatant 1. C) ANS fluorescence of supernatant 2. D) ANS fluorescence of the samples with dissolved pellet.

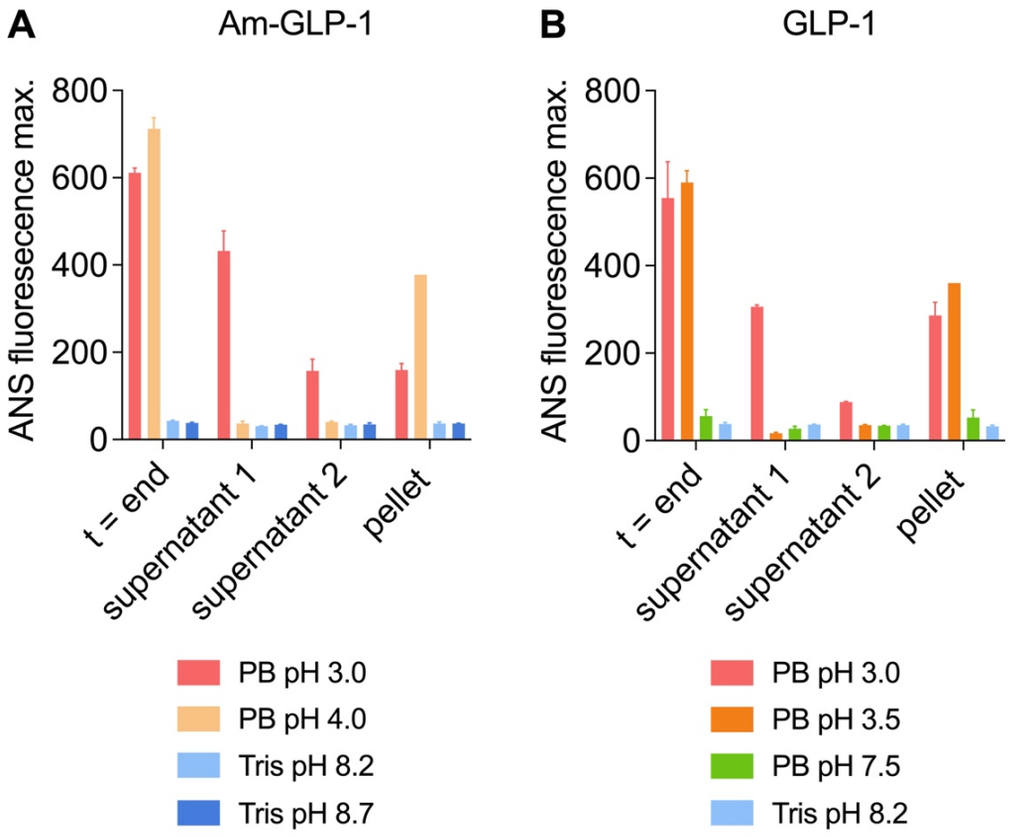


Figure 4.25. ANS fluorescence experiments to determine the specificity of ANS binding: maximum fluorescence signal.

Maximum ANS fluorescence of 150 µM Am-GLP-1 or GLP-1 samples after 140 h of incubation at 37 °C in a plate reader, as well as of the supernatant 1, supernatant 2 and the pellet dispersed in buffer. A) Am-GLP-1 incubated in 25 mM citrate buffer at pH 3.0, citrate buffer pH 4.0, Tris buffer at pH 8.2 (8.5 at 25 °C) and Tris buffer at pH 8.7 (9.0 at 25 °C). B) GLP-1 incubated in citrate buffer at pH 3.0, citrate buffer at pH 3.5, phosphate buffer at pH 7.5 and Tris buffer at pH 8.2 (8.5 at 25 °C).

It was anticipated that the centrifugation steps should effectively separate the fibrils from soluble oligomers and monomer. The ANS, ThT and tryptophan fluorescence were measured for supernatant 1, supernatant 2 and the final resuspended pellet sample.

4.11.1 ANS Fluorescence: Higher Under Acidic Conditions

The ANS data, Figure 4.24A, clearly shows that for both GLP-1 and Am-GLP-1 after aggregation, there are only significant levels of ANS fluorescence at acidic pH values. At neutral or basic pHs, the ANS fluorescence is low, Figure 4.24A. At pH 3.0, the first and second supernatants after centrifugation still show 50-70% of the ANS fluorescence in comparison to sample measured

before any separation step, Figure 4.24B & C. This suggests that either the centrifugation is insufficient to separate monomers/oligomers from fibrils, or that the fibrils also bind to ANS and increase its fluorescence. GLP-1 at pH 3.5 and Am-GLP-1 at pH 4.0 show no ANS fluorescence in the supernatants but fluorescence in the final resuspended pellet sample. This indicates that ANS may well bind to fibrils, increasing its fluorescence, Figure 4.24D. Alternatively, this could also be caused by any oligomeric aggregates binding to species that are pelleted during the centrifugation step. A plot of the maximum ANS fluorescence for all samples and measurements is shown in Figure 4.25A & B.

4.11.2 Comparison to Tryptophan Fluorescence and Change of λ_{max}

The strength of the tryptophan fluorescence can be used as an estimate of the amount of peptide in solution. Similar to the ANS data, the tryptophan fluorescence is stronger at acidic pH values, Figure 4.26A. The peaks at neutral or basic conditions are approx. 3-4 times lower than in acidic conditions, Figure 4.26A. Again, the data shows a high fluorescence signal in supernatant 1 at pH 3.0, Figure 4.26B, indicating that the centrifugation step either failed to pellet all of the aggregated material, or that not all the peptide had formed insoluble aggregates. Similarly, there is a clear peak at pH 8.2, Figure 4.26B. This shows that for GLP-1 at pH 8.2, large aggregates/fibrillar species were not completely separated by the centrifugation step, or that there was still a measurable equilibrium between soluble and insoluble species under these conditions.

In addition, the fluorescence of tryptophan, particularly the λ_{max} value, is known to be influenced by its solvent exposure and degree to which it is buried within a hydrophobic environment. Therefore, it can be used as a probe of the environment of the tryptophan side chain in the different aggregated states. Following Vivian and Callis (2001), the wavelength of emission of Trp is primarily determined by the electrical potential difference across the long axis of the tryptophan's indole ring. For example, positive charges create a redshift when acting on the benzene ring end and a blue shift when acting on the pyrrole ring. For negative charges, the reverse is true.

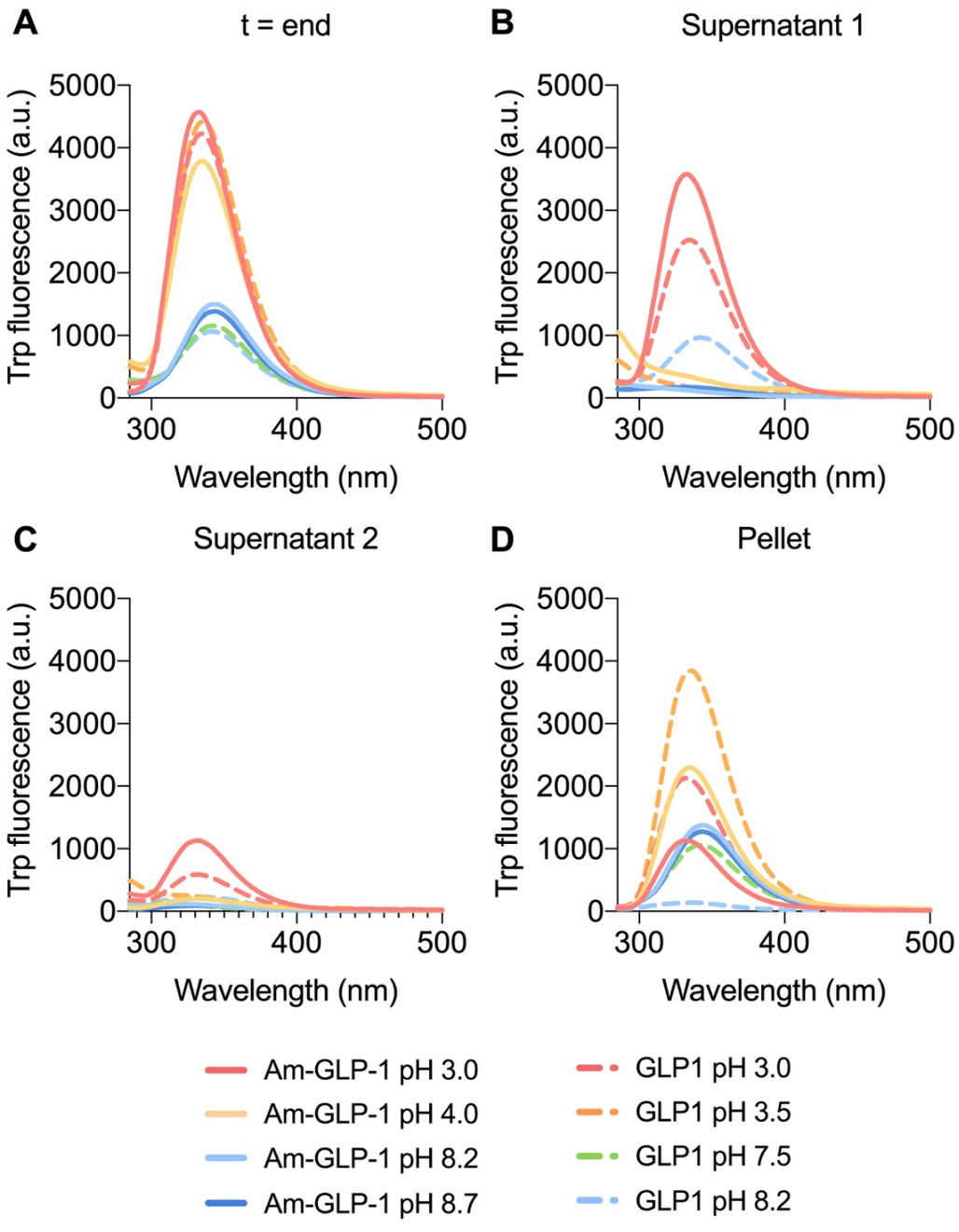


Figure 4.26. Tryptophan fluorescence experiments: after aggregation and centrifugation steps.

Tryptophan fluorescence assay of 150 µM Am-GLP-1 and GLP-1 samples. Am-GLP-1 incubated at 37 °C in citrate buffer at pH 3.0, citrate buffer pH 4.0, Tris buffer at pH 8.2 (8.5 at 25 °C) and Tris buffer at pH 8.7 (9.0 at 25 °C). GLP-1 incubated at 37 °C in citrate buffer at pH 3.0, citrate buffer at pH 3.5, phosphate buffer at pH 7.5 and Tris buffer at pH 8.2 (8.5 at 25 °C). The used buffer concentration was 25 mM. A) Tryptophan fluorescence after 140 h of incubation at 37°C in plate reader B) Tryptophan fluorescence of supernatant 1. C) Tryptophan fluorescence of supernatant 2. D) Tryptophan fluorescence of the samples with resuspended pellet.

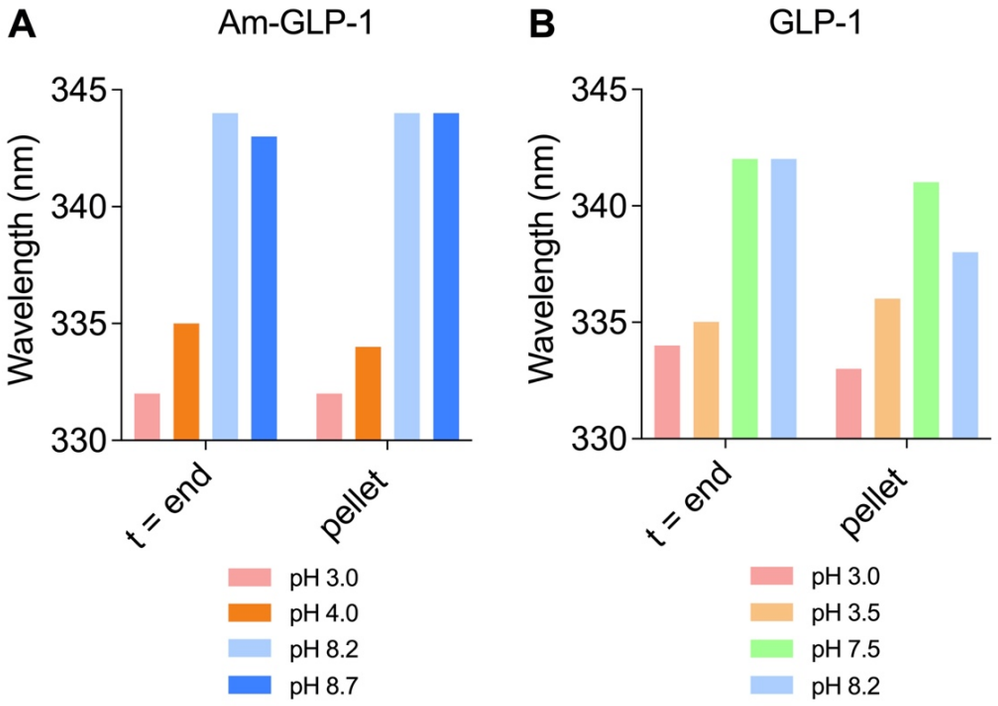


Figure 4.27. Tryptophan fluorescence experiments: λ_{max}.

Tryptophan λ_{max} of 150 µM Am-GLP-1 and GLP-1 samples after 140 h of incubation at 37 °C and in 25 mM buffer in a plate reader before and after two centrifugation and separation steps. A) Am-GLP-1 in citrate buffer at pH 3.0, citrate buffer pH 4.0, Tris buffer at pH 8.2 (8.5 at 25 °C) and Tris buffer at pH 8.7 (9.0 at 25 °C). B) GLP-1 in citrate buffer at pH 3.0, citrate buffer at pH 3.5, phosphate buffer at pH 7.5 and Tris buffer at pH 8.2 (8.5 at 25 °C).

The influence of water is, in most cases, difficult to anticipate, but if one or both faces of the benzene ring are water exposed, the wavelength of the fluorescence peak is more likely close to 350 nm and therefore red-shifted as is seen for monomeric samples of Am-GLP-1 and GLP-1. The λ_{max} of both peptides, GLP-1 and Am-GLP-1, was also determined after the aggregation assay, Figure 4.27A. Both peptides show a pH dependence of λ_{max}. After aggregation, the maximum wavelength is around 335-340 nm under acidic conditions (pH 3.0 to 4.0) and 340-345 nm under neutral or basic conditions. The tryptophan fluorescence of the resuspended pellet (Figure 4.27B) shows a similar pH dependence of λ_{max} values compared to the measurements of Am-GLP-1 and GLP-1 directly after the aggregation assay and without the centrifugation step, Figure 4.23. Therefore, it is most likely that the structure of the samples is conserved during the centrifugation steps.

4.11.3 Separated Aggregates Interact with ANS and ThT

Additional information about the formation of fibrils comes from *ex-situ* ThT experiments. As expected, there is no general difference in the intensity of the ThT fluorescence signal between the acidic and basic samples, Figure 4.28A, as amyloid-like fibrils are formed under both conditions. Interestingly, Am-GLP-1 at pH 3.0 and GLP-1 at pH 3.0 and pH 8.2 both show strong ThT fluorescence in supernatant 1, Figure 4.28B, indicating that there must be species containing extensive β-structure in these samples. This may be short fibrils that have not centrifuged down into the pellet or soluble oligomers with sufficient β-structure to bind to ThT.

In contrast, the samples of Am-GLP-1 at pH 4.0, 8.2 and 8.7 as well as GLP-1 at pH 3.5 and 7.5 show significant levels of ThT fluorescence only in the final pellet. From the comparison of ANS with the *ex-situ* ThT fluorescence, ANS shows a low fluorescence gain under acidic conditions. It could be caused by a higher fraction of oligomeric aggregates under neutral and basic conditions, or ANS binding is influenced more by net charge and therefore pH compared to ThT. The behaviour of the ANS fluorescence and its pH dependence is, overall, similar to the intrinsic fluorescence of tryptophan.

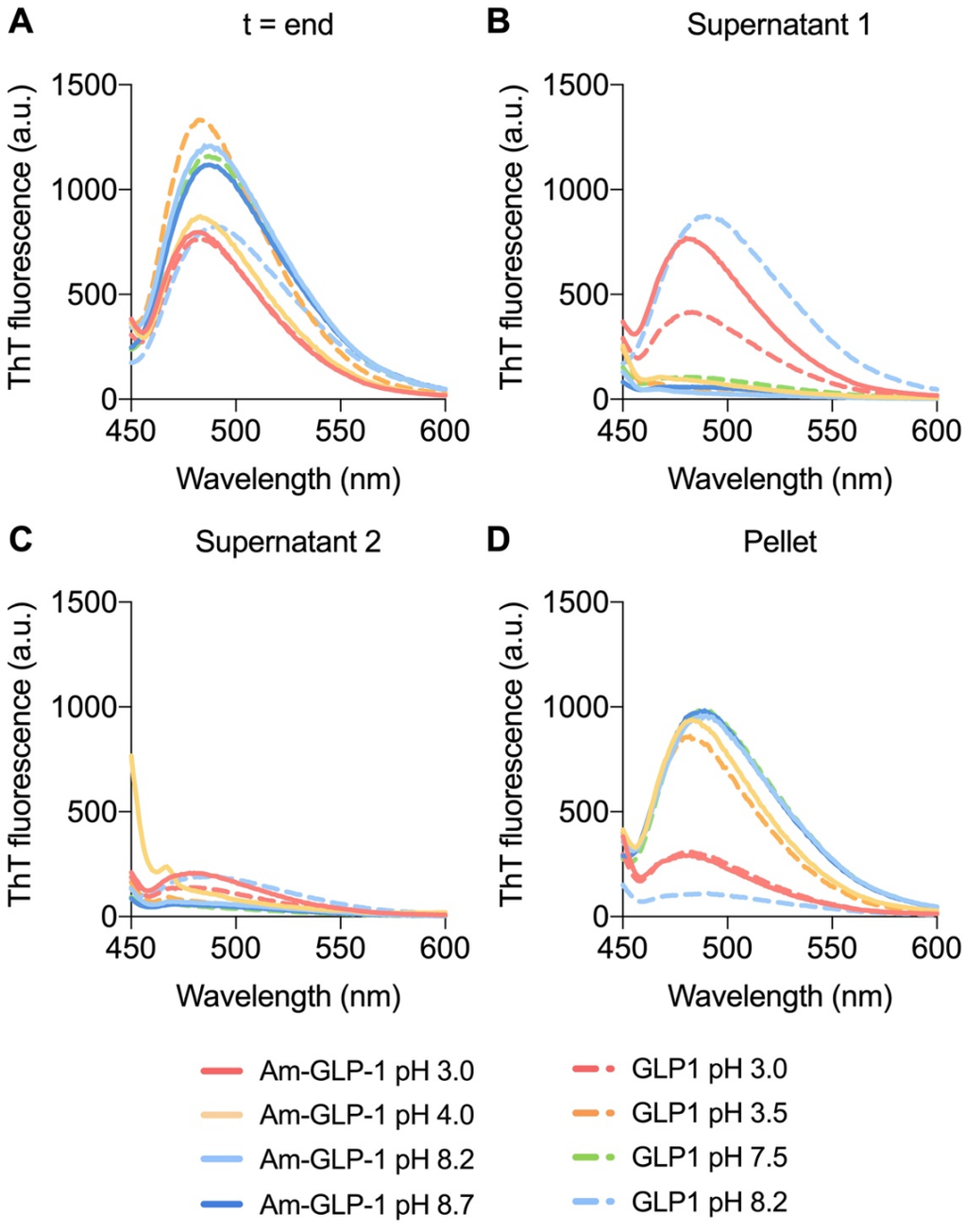


Figure 4.28. *Ex-situ* ThT fluorescence experiments: after aggregation and centrifugation.

Ex-situ ThT fluorescence assay of 150 µM Am-GLP-1 and GLP-1 samples. Am-GLP-1 incubated at 37 °C in citrate buffer at pH 3.0, citrate buffer pH 4.0, Tris buffer at pH 8.2 (8.5 at 25 °C) and Tris buffer at pH 8.7 (9.0 at 25°C). GLP-1 incubated at 37 °C in citrate buffer at pH 3.0, citrate buffer at pH 3.5, phosphate buffer at pH 7.5 and Tris buffer at pH 8.2 (8.5 at 25 °C). All buffers prepared at a concentration of 25 mM. A) *Ex-situ* ThT fluorescence after 140 h of incubation at 37 °C in plate reader B) *Ex-situ* ThT fluorescence of supernatant 1. C) *Ex-situ* ThT fluorescence of supernatant 2. D) *Ex-situ* ThT fluorescence of the samples with resuspended pellet.

4.12 Influence of the C-terminal Amidation on Secondary Structure

The fluorescence data presented in Section 4.10 & 4.11 shows that the hydrophobicity of the aggregates as well as the environment around the tryptophan side chain in the aggregates, most likely fibrils, changes depending on the pH. To assess whether these changes are also accompanied by changes in secondary structure, far-UV CD spectroscopy was undertaken on both peptides both before and after aggregation over a range of pH values.

In Figures 4.29A, 4.30A and 4.31A, the far-UV CD spectra of GLP-1 and Am-GLP-1 are shown before aggregation under acidic, close to neutral and basic conditions. While the far-UV CD spectra were recorded, the HT voltage was always kept below 500 V to guarantee that enough photons are being sampled. At pH values between 3 and 4, both peptides have the same secondary structure, but this structure differs with pH, Figure 4.29A.

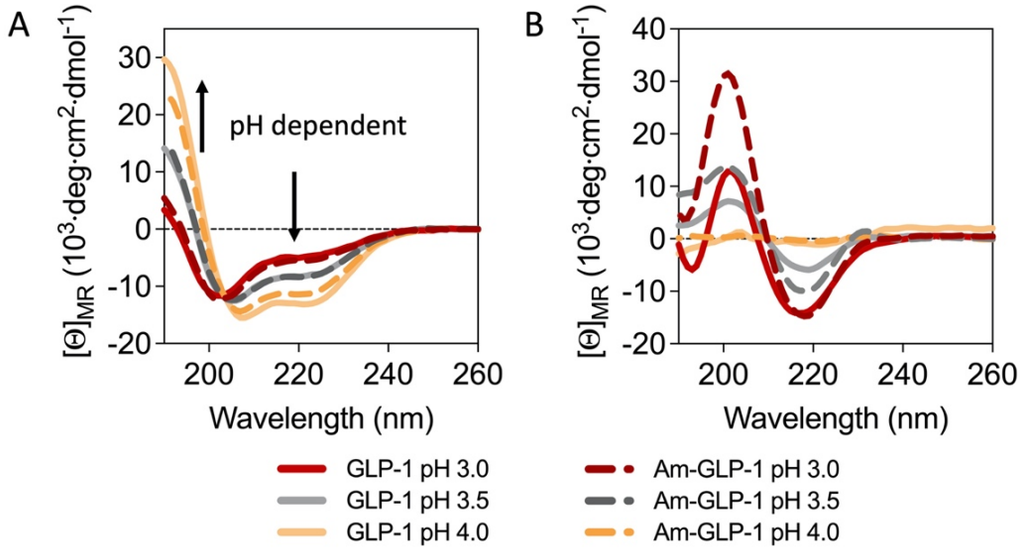


Figure 4.29. Secondary structure of GLP-1 and Am-GLP-1 at acidic pH before and after aggregation.

A) Far-UV CD spectra for GLP-1 and Am-GLP-1 in 25 mM citrate buffer at pH 3.0, 3.5 and 4.0 before aggregation. B) Far-UV CD spectra of the same samples as in Figure 4.29A after 144 h aggregation in the plate reader under the same conditions as the ThT assay.

There is a significant increase in α-helical structure and a decrease of all other structures on increasing the pH from 3.0 to 4.0. At pH values close to neutral, both peptides have the same secondary structure and this does not change between pH 5.0 and 7.0, Figure 4.30A. At basic pH values, there is a significant difference in the far-UV CD spectra between the two peptides but the secondary structure is independent of pH over the range from pH 8.0 to 9.0 (At 37 °C, during the incubation, pH values of the Tris buffer samples change to pH 7.7, 8.2 and 8.7) C), Figure 4.31A. The differences in secondary structure between GLP-1 and Am-GLP-1 observed at higher pH values are likely due to the difference in net charge under these conditions as a consequence of C-terminal amidation. At pH 8.5, the net charge is estimated to be -1.2 for GLP-1 and -0.2 for Am-GLP-1, Figure 4.3B (a change of 600%). In comparison, it is interesting to note that at pH 3.0 the net charge of GLP-1 and Am-GLP-1 are estimated to be 3.8 and 4.7, respectively. Here, the net charge of Am-GLP-1 is only 24% higher than for GLP-1. This increases to a relative change of 45% at pH 4.0 as the net charge of GLP-1 and Am-GLP-1 is calculated to be 2.2 and 3.2, respectively, under these conditions.

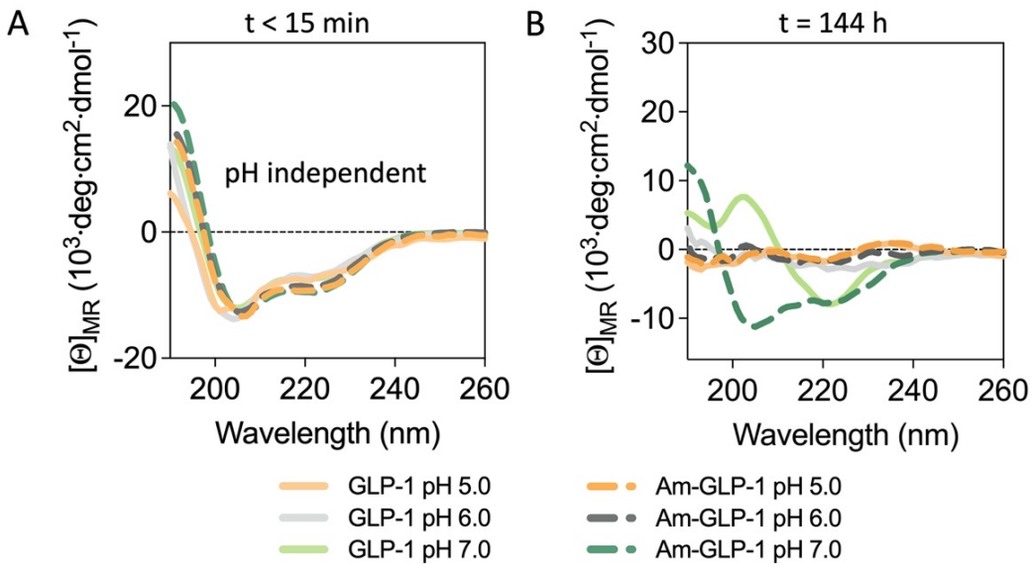


Figure 4.30. Secondary structure of GLP-1 and Am-GLP-1 at acidic and neutral pH before and after aggregation.

A) Far-UV CD spectra for GLP-1 and Am-GLP-1 in 25 mM citrate buffer at pH 5.0 and 6.0 as well as in 25 mM phosphate buffer at pH 7.0 before aggregation. B) Same samples as in Figure 4.29C after 144 h of aggregation in the plate reader.

After aggregation, there is a large change in the far-UV CD spectra of both peptides at most pH values indicative of β-sheet formation (maxima and minima at ~202 and ~218 nm), Figures 4.29B, 4.30B and 4.31B. There are different intensities in signal between the samples, with the signal being particularly low at pH 4.0-6.0. This is attributed to inaccuracy in determining the peptide concentration due to low concentrations resulting from precipitation. In these cases, the peptide concentration has probably been overestimated because of light scattering in the aggregated sample.

In the basic regime, differences between the secondary structures of the aggregates formed by GLP-1 and Am-GLP-1 are observed, as well as changes in maxima/minima with pH, Figure 4.31B. Although overall the aggregates look like they have largely β-sheet structure, there is evidence that the secondary conformation within the fibrils differs.

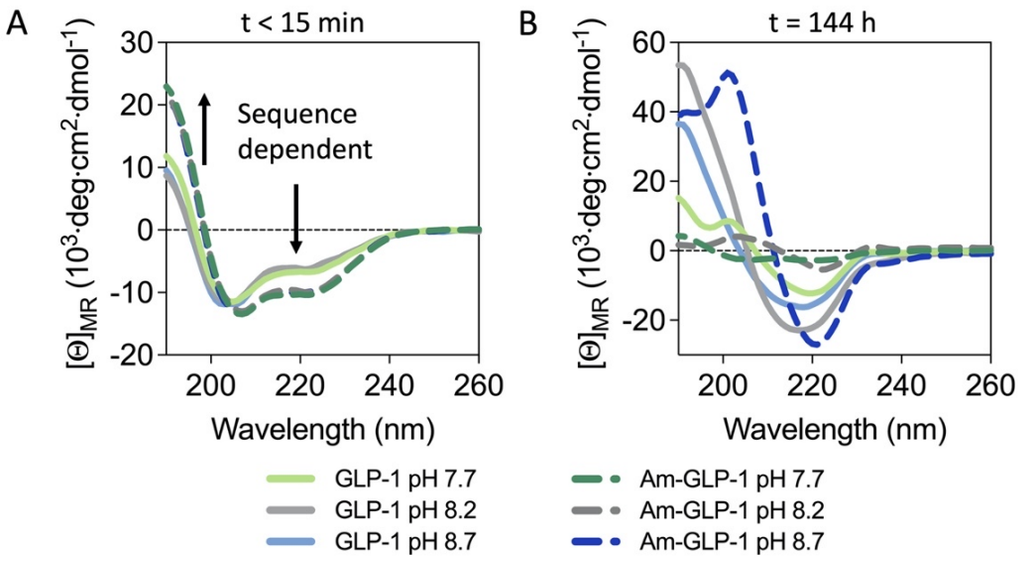


Figure 4.31. Secondary structure of GLP-1 and Am-GLP-1 at neutral and basic pH before and after aggregation.

A) Far-UV CD spectra for GLP-1 and Am-GLP-1 in 25 mM Tris buffer at pH 8.0, 8.5 and 9.0 before aggregation and at 25 °C. B) Far-UV CD of the same samples as in Figure 4.29E after 144 h aggregation at pH 7.7, 8.2 or 8.7 in the plate reader under the same conditions as the ThT assay. The far-UV CD spectra were recorded at 25 °C, whilst the aggregation assays were performed at 37 °C. Thus, the pH values in the Tris buffers were 0.3 pH units higher when the far-UV CD spectra were acquired compared to the pH at which aggregation took place in the plate reader.

The data were analysed using CDPro, which estimates the different fractions of regular α-helix, distorted α-helix, regular β-sheet, distorted β-sheet, turns and disordered structures.[327,328,359] It is worth noting that the analysis with CDPro is based on previously gained structural data from peptides and proteins. Therefore, the analysis always has to be taken to a certain degree qualitatively. However, CDPro still enables a comparison between different sets of far-UV CD data of peptides collected in a systematic manner. For reasons of simplicity, the regular and distorted fractions of either α-helix or β-sheet were combined into one single value of α-helix and β-sheet, respectively. Figure 4.32A & C clearly show how the different elements of secondary structure change with pH for the monomeric peptides before aggregation. There is an increase in α-helicity and concomitant decrease in β-sheet for both peptides between pH 3.0 and 4.5. Overall, especially between pH 6.0 and 9.0 (25 °C), Am-GLP-1 has a slightly higher propensity to form helical structure than GLP-1, Figure 4.32A. These data also show that at pH 3.0, both, Am-GLP-1 and GLP-1 form more β-sheet than α-helical structure within the largely monomeric solution (the system is referred to as being in **State B**). Am-GLP-1 shows the same behaviour additionally at pH 8.5 and 9.0 (25 °C). All other monomeric solutions show higher fractions of α-helix (the system is referred to as being in **State A**). Figure 4.32B & D show the different amounts of secondary structure after aggregation as a function of pH. A clear increase in β-sheet structure can be seen for both peptides over most pH values as expected for amyloid-like fibril formation. The raw data used for the CDPro analysis were also plotted as 3D surfaces, Figures 4.33 and 4.34. The main difference between the Am-GLP-1 and GLP-1 secondary structure data can be seen as the blue embayment around 225 nm and at neutral pH, which is only present for the amidated version of the peptide, Figure 4.33A & B.

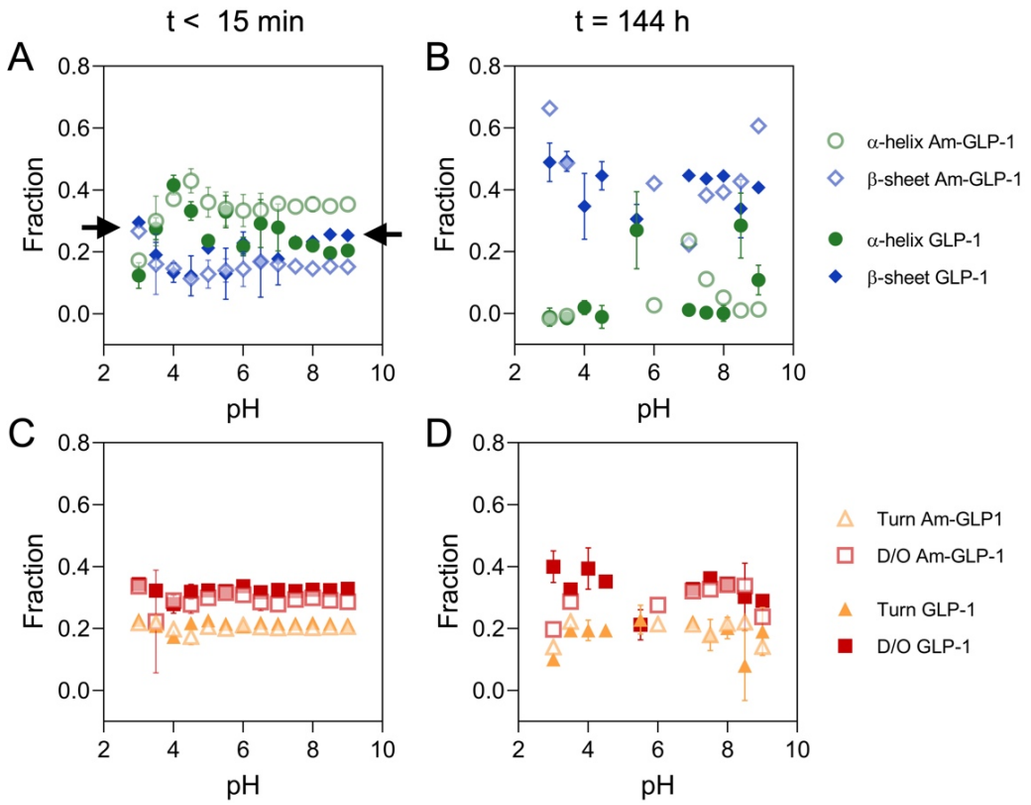


Figure 4.32. Fractions of secondary structure of GLP-1 and Am-GLP-1 before and after aggregation as a function of pH.

The fractions are calculated from the far-UV CD spectra using CDPro. A) Fraction of α-helical and β-sheet structure of Am-GLP-1 and GLP-1 after the peptides were dissolved, filtered and measured within 15 min. The conditions under which either Am-GLP-1 or GLP-1 and Am-GLP-1 show a higher fraction of β-sheet are marked with a black arrow (**State B**). All other conditions show higher fractions of α-helical structure (**State A**) B) Fraction of α-helical and β-sheet structure of Am-GLP-1 and GLP-1 after aggregation. C) Fraction of turn and disordered structure (D/O) of Am-GLP-1 and GLP-1 after the peptides were dissolved, filtered and measured within 15 min. D) Fraction of turn and disordered structure (D/O) of Am-GLP-1 and GLP-1 after aggregation. The errors shown are the standard deviation of the results using three different algorithms to calculate the secondary structure, Section 2.10.

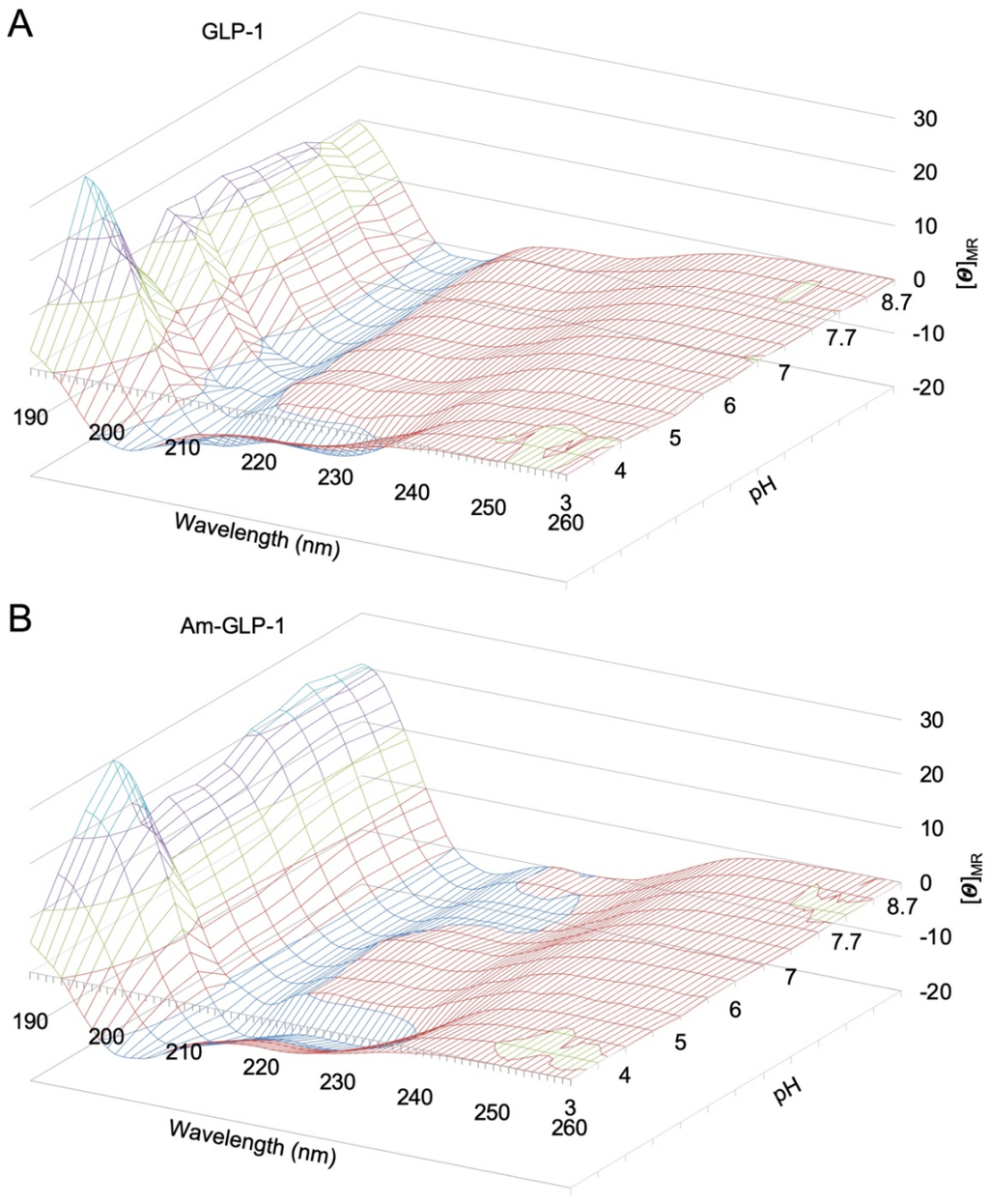


Figure 4.33. Three-dimensional far-UV CD spectral landscape illustrating the differences in secondary structure of GLP-1 and Am-GLP-1 at different pH values using freshly prepared samples.

Far-UV CD spectra recorded directly after the peptides were dissolved in aqueous buffer. Each graph shows the far-UV CD spectra at 13 different pH values from pH 3.0 to 9.0 at 25 °C. The buffers were citrate (pH 3.0 - 6.0), phosphate (pH 6.5 - 7.5) and Tris (pH 8.0 - 9.0, 25 °C) and all the buffer concentrations were 25 mM. The far-UV CD spectra were recorded at 25 °C, whilst the aggregation assays were performed at 37 °C. Thus, the pH values in the Tris buffer were 0.3 °C pH units higher when the far-UV CD spectra were acquired compared to the pH values of 7.7, 8.2 and 8.7 at which aggregation took place in the plate reader. All measurements were done at a peptide concentration of approx. 150 µM or slightly below if the solubility was low. A) Secondary structure of GLP-1 B) Secondary structure of Am-GLP-1.

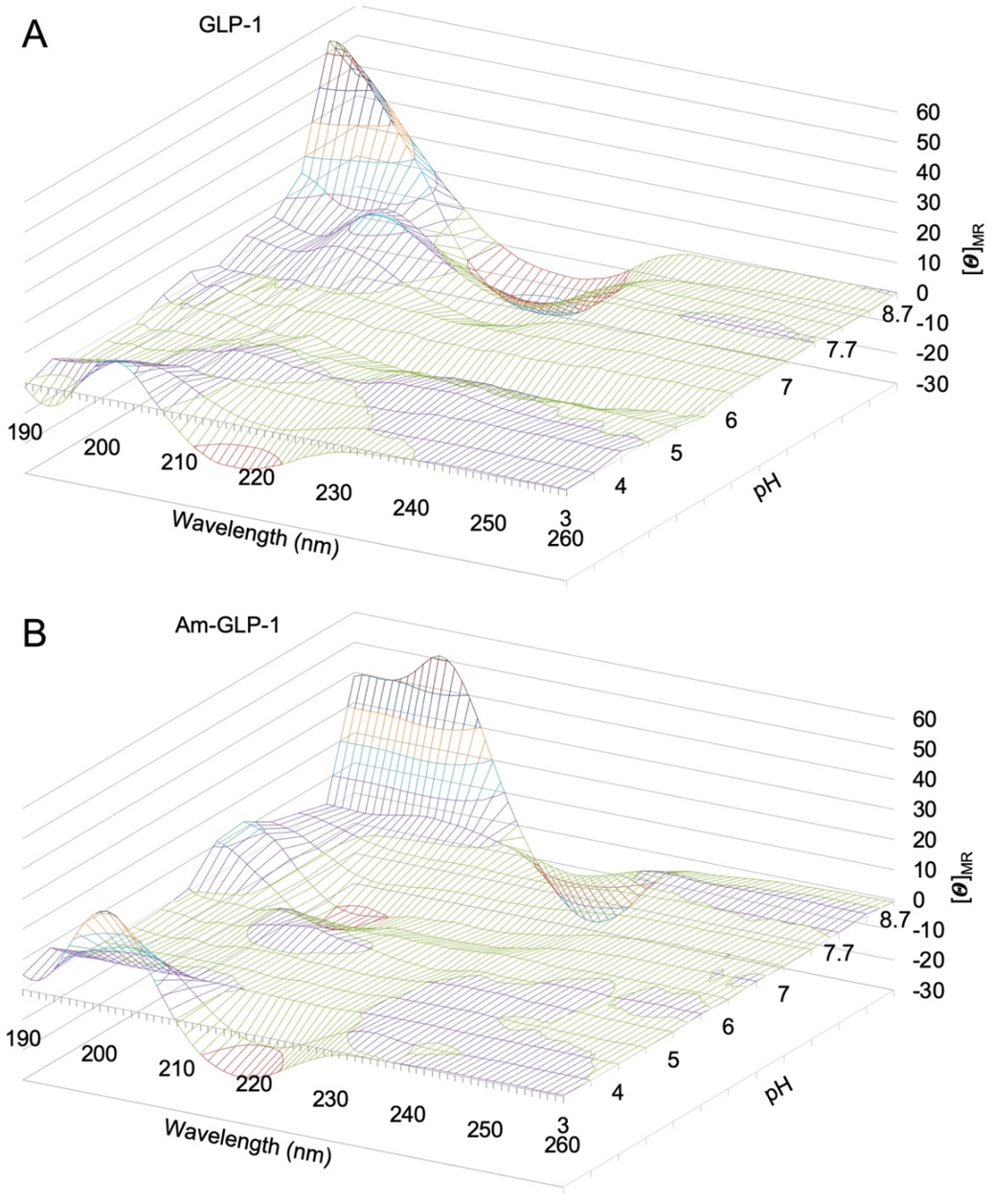


Figure 4.34. Three-dimensional far-UV CD spectral landscape illustrating the differences in secondary structure of GLP-1 and Am-GLP-1 at different pH values after 144 h of aggregation.

Far-UV CD spectra recorded after the ThT aggregation assay. Each graph shows the far-UV CD spectra of 13 different samples aggregated at pH values from pH 3.0 to 8.7 at 37 °C. The buffers were citrate (pH 3.0 - 6.0), phosphate (pH 6.5 - 7.5) and Tris (pH 7.7 - 8.7) and all buffer concentrations were 25 mM. All measurements were done at a peptide concentration of approx. 150 µM or slightly below if the solubility was low. During the spectroscopic analysis the pH of the Tris buffers used during the aggregation assay changed from pH 7.7 to 8.0, 8.2 to 8.5 and pH 8.7 to 9.0 because of the temperature change from 37 °C (aggregation assay) to room temperature (spectroscopic analysis). A) Far-UV CD spectrum of GLP-1 B) Far-UV CD spectrum of Am-GLP-1.

4.13 Discussion

To assess the effect of amidation of the C-terminus on the aggregation behaviour of the therapeutic peptide GLP-1 measurements on the solubility, charge and aggregation kinetics of Am-GLP-1 were made over a wide range of pH values. In addition, ANS assays, tryptophan assays and far-UV CD measurements, characterising the hydrophobicity, burial of tryptophan residues, secondary structure both before and after aggregation, were also undertaken. Finally, the nature of species present at the end of the aggregation was assessed using imaging techniques. A number of systematic differences between the behaviour of GLP-1 and Am-GLP-1 were observed and, in some cases, these could be linked to changes in charge state. Collectively, the data provides more information on the mechanism of self-assembly of GLP-1 based peptides, confirms the presence of off-pathway species populated during fibrillation and illustrates how sensitive aggregation reactions are to small changes in the intrinsic properties of the peptide as well as the environment.

4.13.1 The Mechanism of Aggregation: The Role of On- *versus* Off-Pathway Oligomers

In these studies, with GLP-1 and Am-GLP-1, three different types of aggregation behaviour in terms of how the kinetic parameters $t_{1/2}$ and t_{lag} depend on peptide concentration have been observed. These are assigned to one of three classes. **Class I**: lag time and $t_{1/2}$ decrease with increasing peptide concentration. This is what has been observed for many other peptides[85] and is consistent with standard nucleation-elongation models, Figure 1.7A. **Class II**. The lag time and $t_{1/2}$ do not vary significantly with peptide concentration. This is in agreement with either i) a critical micelle concentration (CMC) model in which the peptide monomer concentration is kept approximately constant above a threshold concentration (the CMC) above which excess peptide monomers arrange in off-pathway aggregates, for example micelles, Figure 1.7C.[86,360] Alternatively, it is also consistent with ii) an off-pathway model as shown by Powers & Powers[89] in which, if aggregation is dominated by

elongation and the release of the monomers from the off-pathway species, $t_{1/2}$ does not change with peptide concentration. This is illustrated in Figure 1.7D. **Class III**. The kinetic parameters (lag time and $t_{1/2}$) increase with increasing peptide concentration. This is a particularly interesting case and rather unusual behaviour which suggests that not only are off-pathway species formed but that the off-pathway species actively interferes with the formation of nuclei or fibrils, Figure 1.7E.[347] This behaviour was observed at pH 3.5 and 7.5 for GLP-1 and pH 8.2 for Am-GLP-1, Figure 4.6. It is also known that a lipidated analogue of GLP-1 (Liraglutide)[91] as well as several other peptides (single-chain Aβ dimer or lysozyme)[347] show this unusual behaviour.

Fitting of ThT data to an NEF model, generates values for δ, a measure of the rate of primary nucleation and κ, a measure of the rate of secondary nucleation. δ varies tenfold across pH values for both GLP-1 and Am-GLP-1, but they vary differentially. From these data, it is clear that the primary nucleation rate increases with increasing peptide concentration under conditions showing typical nucleation-polymerisation behaviour, for example, at pH 3.0 for both peptides. However, For GLP-1 at pH 3.5 and 7.5 (conditions where off-pathway oligomers are thought to be populated), it decreases with increasing peptide concentration, suggesting that off-pathway events are happening. Under the same conditions, κ also decreases with increasing peptide concentration which is the expected behaviour if off-pathway oligomers not only depopulate the monomer concentration but actively suppress secondary nucleation.

Previously, the behaviour of increasing lag time with increasing peptide concentration was attributed to the formation of off-pathway species alone along with a unimolecular step *en route* to fibril formation.[5] However, if both pathways are independent of each other and don`t interfere actively during the aggregation process, this model would result in t_{lag} remaining constant and only $t_{1/2}$ would increase with increasing peptide concentration as has been shown by Powers & Powers in their work on off-pathway aggregation[89], Figure 1.7D. A shorter mathematical approach to prove that an off-pathway system as described in Figure 1.7D cannot lead to longer lag times with increasing peptide concentration is shown in Section 4.8. There, it is mathematically shown that

the amount of available monomer concentration available for the on-pathway route, if the on-pathway is independent of the off-pathway species, cannot be smaller at any time point as the monomer concentration of a sample starting with a lower total monomer concentration at the same time point. It always results in a more rapid formation of both, on- and off-pathway aggregates, and can only approach a constant lag time if the off-pathway acts as a sink for free monomer, Section 1.5.2.

4.13.2 Role of Net Charge on Solubility and Aggregation Propensity

One important factor influencing the interaction of monomers, oligomers, or fibrils with the solution or each other is the net charge of the peptide. Therefore, net charge on the peptide has the potential to influence the aggregation kinetics and even completely change the aggregation pathway.

Amidation of GLP-1 leads to a change in the net charge over pH values between pH 3 and 9 and results in the peptide having a particularly wide-ranging minimum in charge around its pI value, Figure 4.3B. This can be explained by the distance from the pI to the pK_a of the next ionisable group. The pI of Am-GLP-1 is shifted to a value which is further away from the pK_a values of other ionisable groups such that the net charge stays close to zero over a wider pH range compared to GLP-1. The data shows the expected dependence between net charge and solubility, Figure 4.3A & B, but also establishes that there is a correlation between net charge and aggregation propensity as measured by $t_{1/2}$ and lag time, Figure 4.35A & B. This is discussed further in the next section.

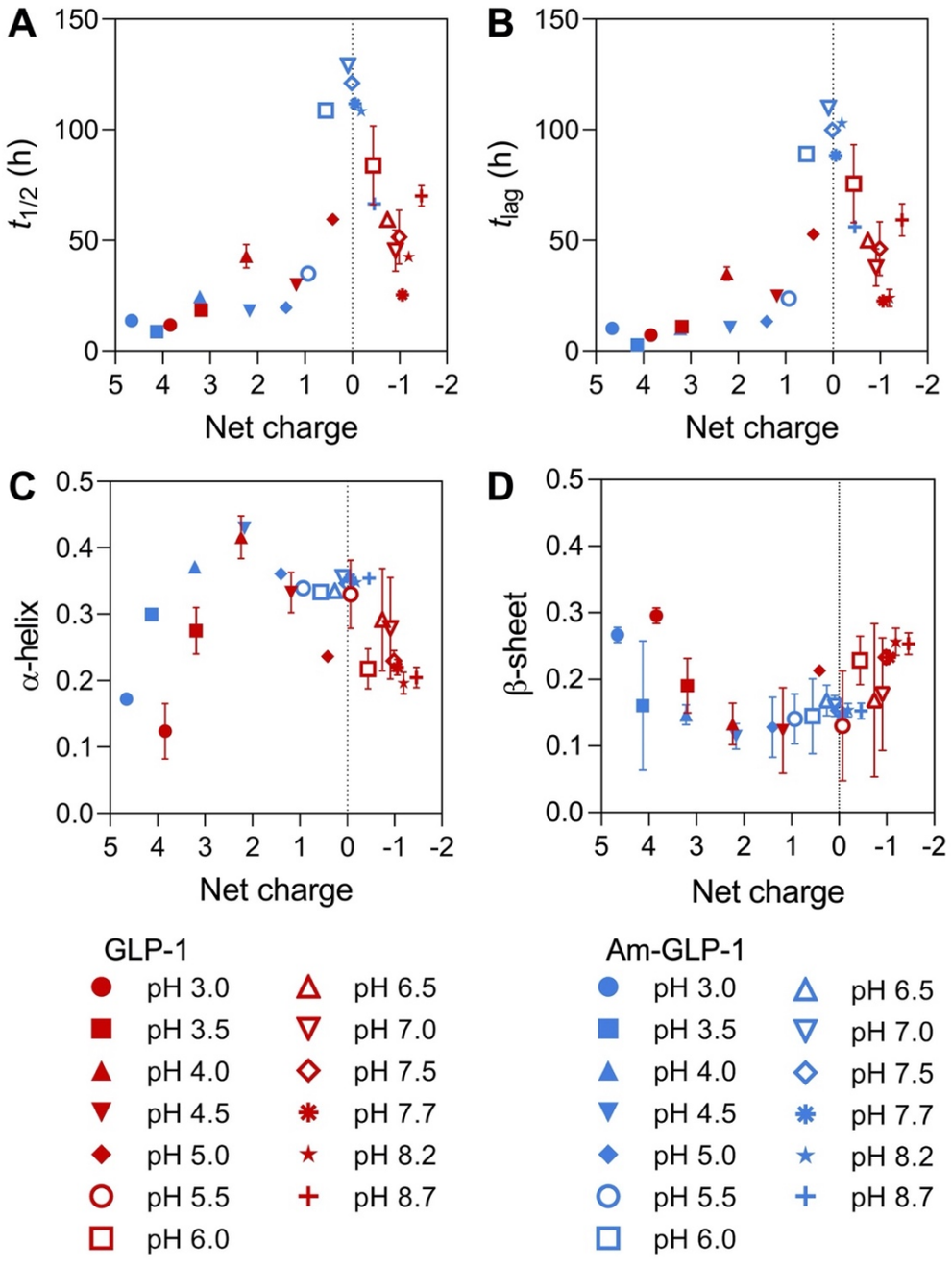


Figure 4.35. Correlation of net charge with aggregation propensity and secondary structure for GLP-1 and Am-GLP-1.

The net charge was calculated as described in Section 2.15. The error bars shown are the standard deviation of the results of three different algorithms used to calculate the secondary structure, Section 2.10. A) $t_{1/2}$ values of GLP-1 and Am-GLP-1 aggregated at pH values between pH 3.0 and 8.7 plotted *versus* the net charge. B) t_{lag} values of GLP-1 and Am-GLP-1 aggregated at pH values between pH 3.0 and 8.7 plotted *versus* the net charge. All measurements were done at a peptide concentration of approx. 150 µM or slightly below if the solubility was low. C) The fraction of α-helical-structure in the monomeric peptide as calculated from far-UV CD spectra (Figure 4.33) is shown *versus* net charge. D) The fraction of β-sheet structure in the monomeric peptide as calculated from the far-UV CD spectra (Figure 4.33) *versus* net charge.

4.13.3 Less Electrostatic Repulsion Does Not Lead to Higher Aggregation Propensity

The results on the aggregation of Am-GLP-1 and GLP-1, Figures 4.4A and B and 4.5, give a broad overview of the aggregation propensity of both peptides at different pH values as well as at different peptide concentrations. It is clear that net charge strongly influences aggregation affecting the kinetic parameters t_{lag} and $t_{1/2}$, Figure 4.35. Somewhat surprisingly, the results clearly show that both peptides have a lower aggregation propensity at pH values close to their pIs. This is in contrast to some models that predict that decreasing the net charge on a peptide increases its aggregation propensity as a result of decreasing the otherwise unfavourable electrostatic interactions that have to be overcome in order to achieve self-assembly.[132–136] From earlier studies on GLP-1, this unusual kinetic behaviour was attributed to the formation of off-pathway species during the aggregation of this family of peptides under specific conditions.[5] The off-pathway oligomers are populated close to the pI of the peptide and act as an off-pathway sink of monomer thus reducing aggregation rates.

4.13.4 Specificity of ANS Fluorescence

One of the methods in this project used to gain more information about peptide oligomers and the structure of peptide aggregates in general is ANS fluorescence, which is widely used to probe the existence of partially structured folding intermediates.[356–358] However, the data shown here suggest that ANS fluorescence is not always highly specific to such intermediates. The data shows either i) that aggregates that interact with ANS cannot be separated from ThT-binding fibrils by centrifugation, or ii) the fibrils themselves have some affinity for ANS and can increase its fluorescence. Thus, care must be taken in interpreting the results of ANS experiments and one should not assume that ANS fluorescence is solely due to its binding to intermediate states/oligomers.

The data also demonstrates that the ANS-assay is strongly affected by pH showing a clear difference in behaviour between acidic and basic pH values. Such a behaviour could be explained by either i) two different forms of aggregates being formed in acidic and basic conditions with different

propensities to bind to ANS, or ii) different binding affinities of ANS to similar species under acidic and basic conditions due to the influence of net charge. Under basic conditions, the ANS binding is so low that the assay is not able to prove whether in the case of GLP-1 at pH 8.2, a significant amount of peptide was in an oligomeric or fibrillar form in the different samples.

If the washed pellets consist only of fibrils which bind to ThT and the fibrils are not disturbed during the centrifugation or washing step, then the results suggest that ANS also binds to fibrils as well as oligomers. However, it cannot completely be excluded that some of the larger oligomers also pellet with the fibril. If this is the case, then the ANS fluorescence might still be due to oligomer binding only.

There is also some evidence, for example at pH 7.5, that ANS can bind under certain conditions specifically to oligomeric intermediates, but such findings have to be taken with care due to the previously described caveats. These experiments all highlight the limitations of using ANS and the care that needs to be taken when interpreting ANS fluorescence data in the context of protein/peptide aggregation.

4.13.5 Structure of the Monomeric Peptides

It is clear that the net charge on the peptide influences both solubility and aggregation propensity. This leads to the question of whether net charge alone is responsible for the difference in aggregation behaviour observed for GLP-1 and Am-GLP-1 or whether the effects are indirect, i.e., the net charge influences other properties of the peptide such as secondary structure that is responsible for the change in aggregation propensity. A number of experiments on freshly prepared samples of GLP-1 and Am-GLP-1 shed light on the structure of the monomeric peptides in solution.

From the ANS and tryptophan fluorescence data, it is likely that the monomer of both peptides studied has little fixed tertiary structure in either acidic or basic conditions, Figure 4.22A-D. In contrast, the secondary structure of the two peptides differs at specific pH values, Figures 4.29, 4.30 and 4.31. The far-UV CD data show that the secondary structure of both peptides is similar at acidic pH values and also shows that both peptides undergo similar pH-dependent

changes in secondary structure in the acidic region. Notably, for both peptides there is an increase of α-helical structure on increasing the pH from 3.0 to 4.0 or 4.5, Figure 4.32A. This is probably caused by protonation of Asp9 or one or more of the three glutamic acid residues at positions 3, 15 and 21, which are likely to have pK_a values around 3.7 and 4.3, respectively.[340] Significantly, there is a correlation between the secondary structure of the monomeric peptides at the start of the aggregation reaction and the net charge, Figure 4.35C & D. Higher net charges favour an extended β-conformation over a more compact α-helical structure presumably because this minimizes electrostatic repulsion between like charges on the same peptide molecule. This, in turn, leads to the question if the amount of α-helical *versus* β-sheet structure in the monomeric peptide and the rate of fibril formation, as shown by $t_{1/2}$ and t_{lag}, is correlated, Figure 4.36.

For both peptides there is no general correlation found between the amount of α-helical and β-sheet structure and the rate of fibril formation, Figures 4.36A & B and 4.37A & B. What can be seen is a difference for Am-GLP-1 under conditions where the peptide is positively charged and negatively charged with the exception of pH 6.0, which is close to the pI. Under conditions at which Am-GLP-1 has a negative net charge all samples show a high α-helical content and longer lag times than for positively charged Am-GLP-1.

The data also shows a relationship between the difference of α-helical and β-sheet structure, Δ, and the rate of aggregation observed under conditions where the peptide follows a nucleation-polymerisation model. For Am-GLP-1 this is at pH 3.0 and for GLP-1, it is at pH 3.0, 8.5 and 8.7, and under most of these conditions there are clear negative Δ values (**State B**), Figures 4.36C & B and 4.37C. Other conditions show either values of Δ close to zero or positive values (**State A**), Figures 4.36C and 4.37C.

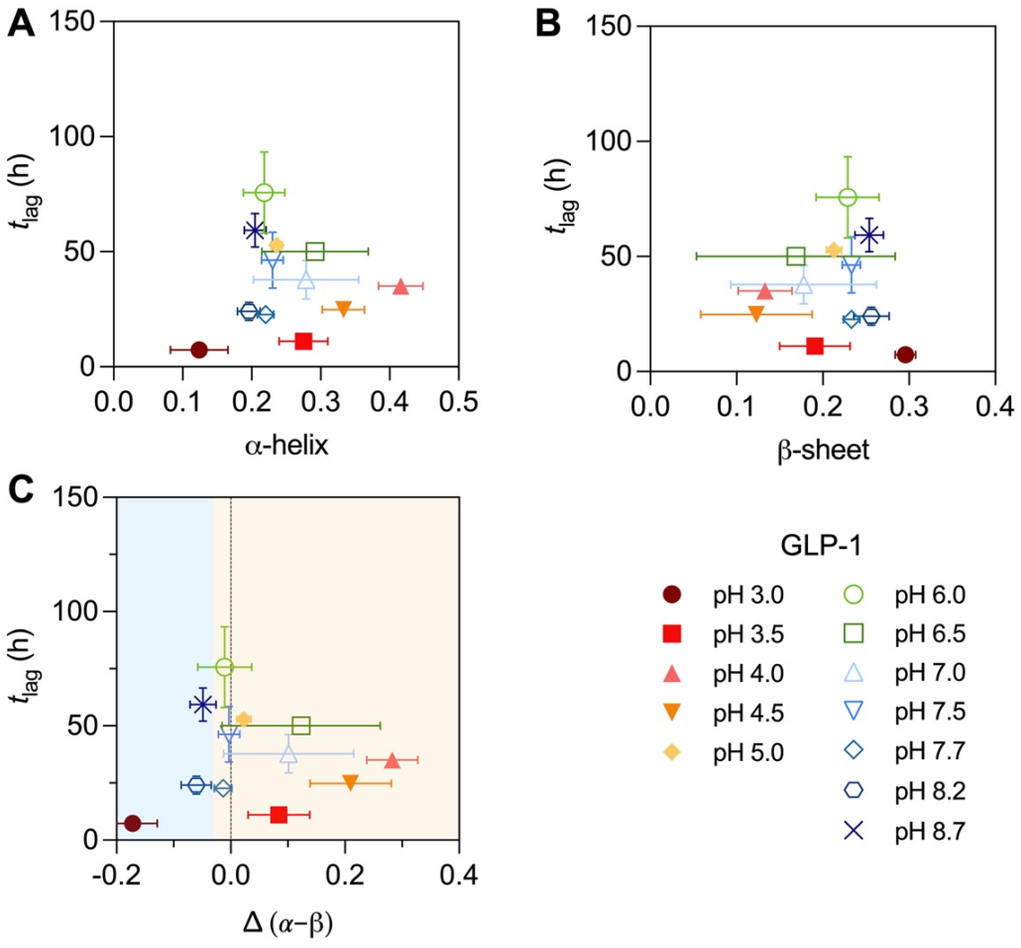


Figure 4.36. Correlation of aggregation propensity (t_{lag} and $t_{1/2}$) and secondary structure for GLP-1.

Data of t_{lag}, determined from data shown in Figure 4.5, plotted *versus* the fraction of α-helical and β-sheet structure, Figure 4.32. The data are shown for GLP-1 aggregated at pH values between 3.0 and 8.7. The far-UV CD spectra used to determine the secondary structure were recorded at 25 °C, whilst the aggregation assays were performed at 37 °C. Thus, the pH values in the Tris buffer were 0.3 °C pH units higher when the far-UV CD spectra were acquired compared to the pH values of 7.7, 8.2 and 8.7 at which aggregation took place in the plate reader. Peptides in samples represented by a solid symbol have a positive net-charge. The error bars shown are the standard deviation of a triplicate measurement of t_{lag} and the results of three different algorithms used to calculate the secondary structure, Section 2.10. A) t_{lag} of GLP-1 plotted *versus* its corresponding fraction of α-helical structure. B) t_{lag} of GLP-1 plotted *versus* its corresponding fraction of β-sheet structure. C) t_{lag} *versus* difference of fraction of α-helical and β-sheet structure Δ of GLP-1. Samples under of the conditions in the blue shaded area (State B) showed constant or increasing t_{lag} with increasing peptide concentration. In the orange shaded region (State A), the samples followed typical nucleation-elongation kinetics.

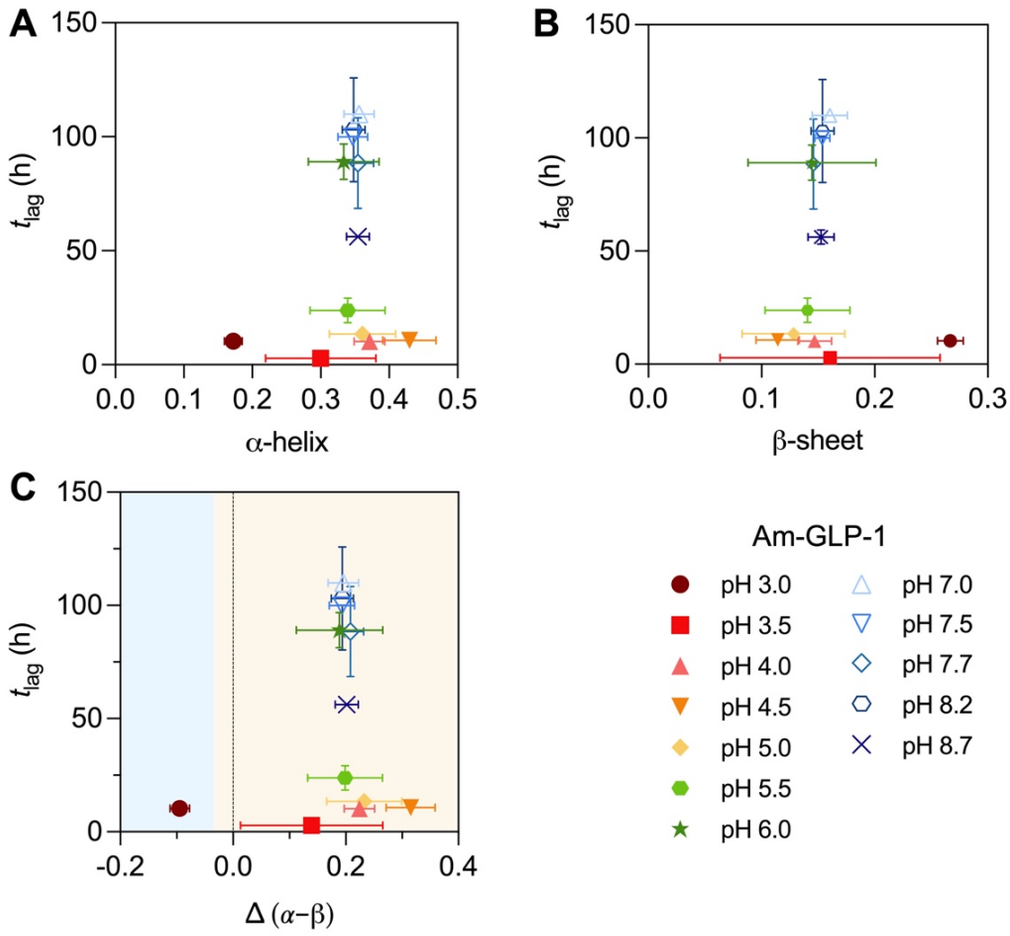


Figure 4.37. Correlation of aggregation propensity (t_{lag} and $t_{1/2}$) and secondary structure for Am-GLP-1.

Data of t_{lag}, determined from data shown in Figure 4.5, plotted *versus* the fraction of α-helical and β-sheet structure, Figure 4.32. The data are shown for Am-GLP-1 aggregated at pH values between 3.0 and 8.7. The far-UV CD spectra used to determine the secondary structure were recorded at 25 °C, whilst the aggregation assays were performed at 37 °C. Thus, the pH values in the Tris buffer were 0.3 °C pH units higher when the far-UV CD spectra were acquired compared to the pH values of 7.7, 8.2 and 8.7 at which aggregation took place in the plate reader. Peptides in samples represented by a solid symbol have a positive net-charge. The error bars shown are the standard deviation of a triplicate measurement of t_{lag} and the results of three different algorithms used to calculate the secondary structure, Section 2.10. A) t_{lag} of Am-GLP-1 plotted *versus* its corresponding fraction of α-helical structure. B) t_{lag} of Am-GLP-1 plotted *versus* its corresponding fraction of β-sheet structure. C) t_{lag} *versus* difference of fraction of α-helical and β-sheet structure Δ of Am-GLP-1. Samples under of the conditions in the blue shaded area (State B) showed constant or increasing t_{lag} with increasing peptide concentration. In the orange shaded region (State A) the samples followed typical nucleation-elongation kinetics.

4.13.6 The Secondary Structure of the Monomeric Peptides Influences Aggregation Propensity and Mechanism

The far-UV CD spectra show that monomers can be found in at least two different conformational states (or ensembles). From CDPro analysis, **State A** is defined by a higher percentage of α-helical than β-sheet structure whilst **State B** has a higher β-sheet content than α-helical structure, Figures 4.32 and 4.38. The α-helical structure is favoured under conditions where the peptide has a low net positive charge whilst higher β-sheet structure is favoured at pH values where the peptide has either a high positive or negative charge, Figure 4.35C & D. In three out of the eight different peptide and pH conditions studied, the peptides were observed in State B. In all of these three cases, a decrease of the lag time and $t_{1/2}$ with increasing peptide concentration was observed in the ThT assays (Class I).

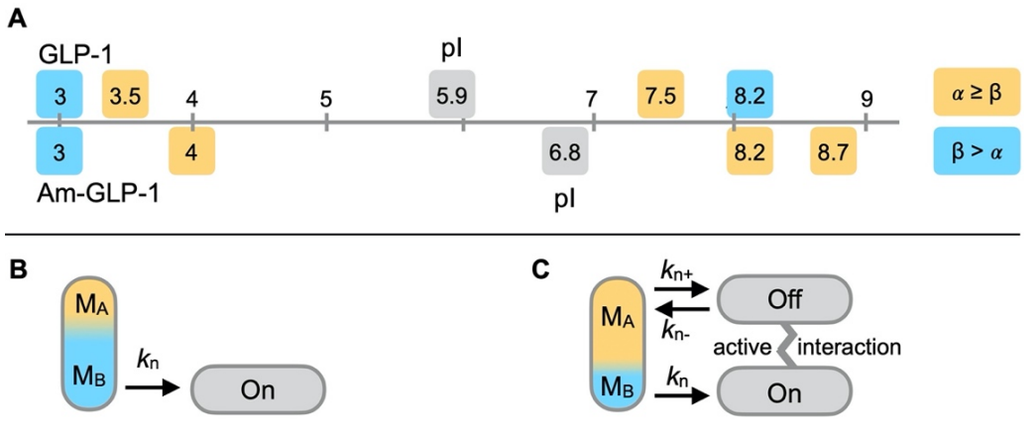


Figure 4.38. Two different monomer conformations (state M_A and M_B) and how these relate to the pI and their role in in forming on- and off-pathway oligomers.

A) Two different conformational states of the monomeric peptide as shown by far-UV CD and their relative predominance at difference pH values. Samples with more or the same amount of α-helical than β-sheet structure are shown in orange (State A) and samples with more β-sheet than α-helical structure are shown in blue (State B). The experimentally determined pI values for GLP-1 and Am-GLP-1 are 5.9 and 6.8, respectively and are shown in grey. B) Conditions in which peptide monomer is in State B follow a predominantly on-pathway mechanism. C) Conditions in which the peptide monomer is in State A have more complex kinetics and show the population of off-pathway species which may or may not actively interact with on-pathway species.

In all the other conditions studied, the peptides were in **State A**. Under these conditions either a constant (Class II) or even increasing (Class III) lag time and $t_{1/2}$ was observed with increasing peptide concentration. The conditions under which **State A** was observed was always closer to the pI in comparison to conditions where the monomeric peptide is in **State B**, Figures 4.35C & D and 4.38.

4.13.7 Structure/Properties of Off-Pathway Oligomers

The observation of different aggregation behaviours for two different monomer ensembles, states A and B, leads to the proposal that monomers with higher α-helical content result in the formation of off-pathway aggregates. These more helical monomers tend to form under conditions where the peptide has a relatively low net charge in comparison to the monomers in state B which have a higher percentage of β-sheet structure and which form at pH values where the peptide has a higher net charge, Figure 4.35. This leads to the conclusion that the net and/or local charge influences the secondary structure of the monomeric peptide, which in turn affects the aggregation behaviour. Thus, the charge state of the peptide influences the propensity of the peptide to aggregate either directly but also likely through induced changes in secondary structure. Based on these results and correlations, it is possible to speculate that the off-pathway oligomers have more α-helical structure and the on-pathway oligomers more β-structure. Although it is not possible to get structural information directly on the oligomeric states, it is interesting to note that all samples which start with higher α-helical content compared to β-sheet are more likely to populate off-pathway oligomers. One might therefore speculate that off-pathway species are more helical than their on-pathway counterparts. Under some conditions, over time α-helical content approaches zero, suggesting that the off-pathway oligomers are converted into β-sheet structures and the amount of the off-pathway oligomers is too small to be distinguished by far-UV CD.

4.13.8 The Structure of Fibrils of GLP-1 and Am-GLP-1 are Sensitive to pH

SEM images at different pH values were recorded to learn more about the structure and morphology of fibrils of Am-GLP-1 formed under different conditions. These images, and previous AFM images of fibrils of GLP-1 at different pH values,[351] show that the higher-order structures of the aggregates formed by GLP-1 and Am-GLP-1 vary significantly with pH. Am-GLP-1 showed different fibril morphologies within the densely covered region of the dried sample, Figures 4.16A & C and 4.17A & C, suggesting that the fibrils formed under different conditions may have distinct surface characteristics which may lead to different higher-order structures. Under acidic conditions, the fibrils tend to stick faster to the glass surface in comparison to fibrils formed under basic conditions. Figure 4.18. This phenomenon could be caused by different surface properties, e.g., hydrophobicity, or simply by different net charge.

More evidence that GLP-1 and Am-GLP-1 fibrils have different surface properties when formed under acidic or basic conditions comes from the ANS and tryptophan fluorescence data. The ANS data shows that fibrils of both peptides formed under acidic conditions bind ANS therefore indicating that there are hydrophobic patches on the surface of the aggregates formed under these conditions. In addition, binding may be facilitated by favourable electrostatic interactions between the positive charge on the peptide and the negatively charged dye at low pH.[361] In contrast, at higher pH no ANS binding is observed. This may be due to the lack of exposed hydrophobic patches formed in the fibrils under these conditions or may simply be a consequence of the lack of favourable electrostatic interactions as the peptide and therefore aggregate has a different net charge.

The results of the tryptophan fluorescence experiments indicate that the Trp side chains in GLP-1 and Am-GLP-1 are buried after aggregation under acidic and basic conditions, but the observed change in λ_{max} is larger under acidic conditions, Figure 4.22E. This could be due to differences in the structure of the fibril formed or differences in the higher-order structures formed under acidic and basic pH. Far-UV CD data supports a model in which the intrinsic

structure of the fibrillar species varies with pH, as despite the fact that there is strong evidence of β-sheet structure in all cases, the far-UV CD spectra of the aggregated species do vary, Figures 4.29B and 4.31B.

4.13.9 Conclusions

Here, we have shown that a C-terminally amidated variant of GLP-1 aggregates to form amyloid-like fibrils over a wide range of condition similar to previous results on GLP-1. The structure, particularly the higher-order structure, of the fibrils formed by Am-GLP-1 varies with pH. However, the pH dependence of aggregation for Am-GLP-1 differs from that observed for GLP-1, which can be attributed to the result of their different pI values and therefore different net charge at different pHs. As with GLP-1, there is strong evidence that Am-GLP-1 also forms off-pathway oligomeric species (as well as on-pathway oligomers that lead to fibril formation) under certain conditions. These off-pathway species, which act to slow but not prevent fibril formation, are populated near to the pI of the peptides, where the net charge on the molecule is close to zero. This results in the opposite effect to that has been reported in earlier studies on other peptides, where low net charge minimises electrostatic repulsion between peptides and thus accelerates the rate of fibril formation. Interestingly, correlations between the net charge on the peptide and the secondary structure it forms in its monomeric state, low positive net charge favouring more α-helical conformations, higher positive or negative net charges favouring more extended β-structures are observed. The consequence is that there is a correlation between the secondary structure in the monomer and its aggregation propensity – more helical structures favouring the formation of off-pathway oligomers resulting in slower aggregation, whilst more β-structures seem to favour the formation of on-pathway oligomers (or disfavour the formation of off-pathway species) and enhance aggregation propensity. Collectively, these studies show how sensitive aggregation reactions are to even small changes in the chemical nature of a peptide as well as to external factors such as pH. In addition, these results suggest that all members of the GLP-1 family of peptides are likely to populate off-pathway oligomers and show a delicate balance between on- and off-pathway processes which can be altered by changing the pH.

5 THE EFFECT OF LIPIDATION ON THE PHYSICAL STABILITY OF PEPTIDES

From the detailed study of GLP-1 and Am-GLP-1, it is clear that C-terminal amidation and the consequent change of net charge has a significant influence on the aggregation characteristics of GLP-1. In this Chapter, the results of studies on the effect of lipidation on the physical stability of GLP-1 are described. Two peptides, called IPP4 and IPP5, were studied. They are lipidated at different positions, the modification involving attachment of a palmitic acid to the amino group of a glutamic acid, which is used as a spacer, the γ-carboxyl group of which is subsequently covalently linked to the ε-amino group of a lysine side chain in the peptide, Figure 5.1.

In the case of IPP5, the lipid-spacer molecule was attached at the first naturally occurring lysine (at position 20) in the Am-GLP-1 sequence. Subsequently, the positive charge of the lysine is lost and an additional negatively charged carboxylate group of the glutamic acid is added. In addition, in IPP5, the lysine at position 28 is substituted by an arginine (K28R) to prevent lipidation at this position. Arginine is chosen as a substitute as it retains the positive charge and shows a similar hydropathicity to lysine. The lipidation and the loss of the positive charge on the lysine side chain on conjugation should increase the

overall hydrophobicity of the peptide, however, the hydrophilicity may be partly replaced by the negative charge of the glutamic acid spacer. IPP5 has the same sequence and modification as Liraglutide but is amidated at the C-terminus.

For IPP4, both naturally occurring lysines are substituted by arginines (K20R and K28R) to prevent covalent modification. In this case, the hydrophilic glutamine at position 17 is replaced with a lysine (Q17K) enabling lipidation at this site, Figure 5.1.

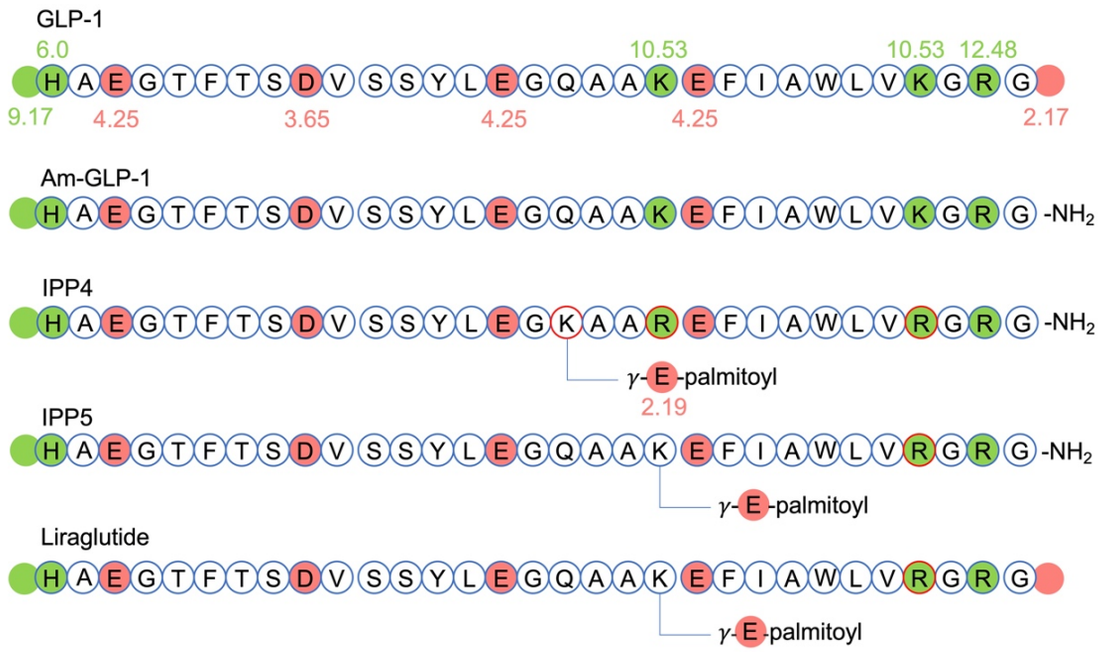


Figure 5.1. Amino-acid sequence of GLP-1, Am-GLP-1, IPP4, IPP5 and Liraglutide.

Possible de-/protonation sites are marked in green or red including the N- and C-terminus of the peptides if not amidated. Green sites are positively charged if the pH is lower than the pK_a of the group (pK_a values are shown in green next to the site for the parent peptide GLP-1). Red sites are negatively charged if the pH is higher than the pK_a value (shown in red next to the site for the parent peptide GLP-1 and the linker group in IPP4). The fraction of de-/protonation and therefore the change of charge around the pK_a can be determined using Equations 2.12 and 2.13.

5.1 Molecular Weight and Purity of IPP4 and IPP5

To measure the molecular weight of the lipidated peptides, as well as to compare the purity of the samples, LC-MS measurements were conducted. The UV absorbance at 210 nm shows a single narrow peak for both peptides, Figures 5.2A and 5.3A. The same peak was found in the total ion count (TIC) data, Figures 5.2B and 5.3B. This establishes that the peptides have high purity and no other molecules with a similar molecular weight which absorb in the UV are present. However, both peptides also show additional smaller peaks at longer elution times. The species corresponding to these peaks do not absorb in the UV and are probably more hydrophobic, causing the longer retention times. Between 5.30 and 6 min, both lipidated peptides also show a broad peak which was not observed for the non-lipidated peptides. These broad peaks are artefacts caused by the LC-MS method used for lipidated peptides and were similarly seen in a run with pure water injected (data not shown).

Figure 5.2C shows the mass spectrum of IPP4 with IPP4 mainly present as a +3 molecular ion (1260.27 m/z). Smaller fractions are also observed as +5 molecular ion (756.85 m/z), +4 molecular ion (945.33 m/z), and +2 molecular ion (1890.12 m/z), Figure 5.4C. The determined molecular weight is 3778.2 ± 0.8 Da, Figure 5.3C. This value is in good agreement with the theoretical value of 3778.26 Da as well as the relative molecular mass of 3778.33 Da obtained by Bachem using matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS).

The mass spectrum of IPP5 shows that it is also mainly present as a +3 molecular ion (1250.85 m/z). Again, some smaller peaks are seen for the +4 molecular ion (938.56 m/z), and +2 molecular ion (1875.83 m/z) but not the +5 molecular ion as the maximum positive charge of IPP5 is +4. The determined molecular weight is 3749.8 ± 0.4 Da, Figure 5.3C. This value is in good agreement with the theoretical value of 3750.2 Da.

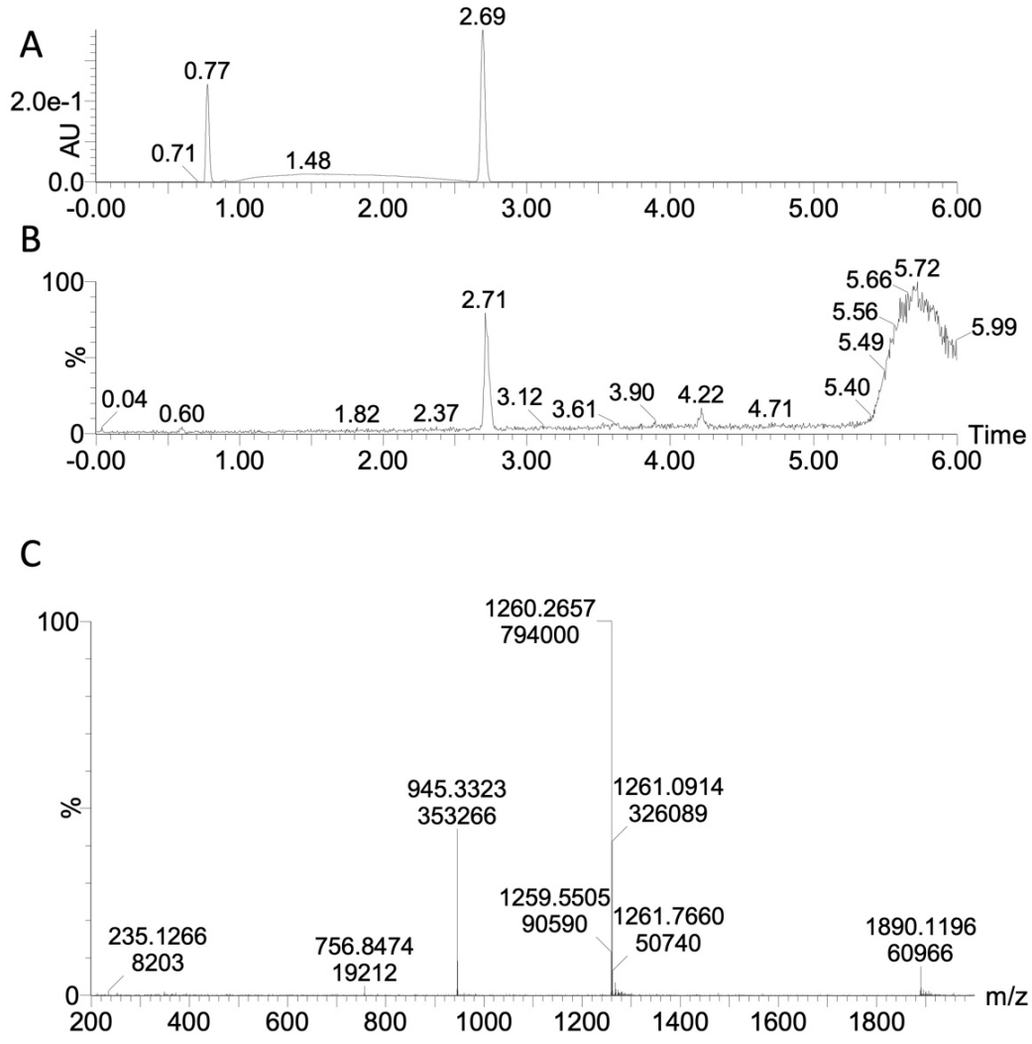


Figure 5.2. LC-MS of IPP4.

A) UV absorbance signal at 210 nm. The peak observed at 0.77 min is caused by the injection of the sample. It is due to the different percentage of acetonitrile in the sample in comparison to the LC gradient as well as the presence of salt in the sample. The x-axis is given in minutes. B) Total ion chromatogram (TIC). The x-axis is given in minutes. C) Mass spectrum of the peptide peak eluting at 2.71 min.

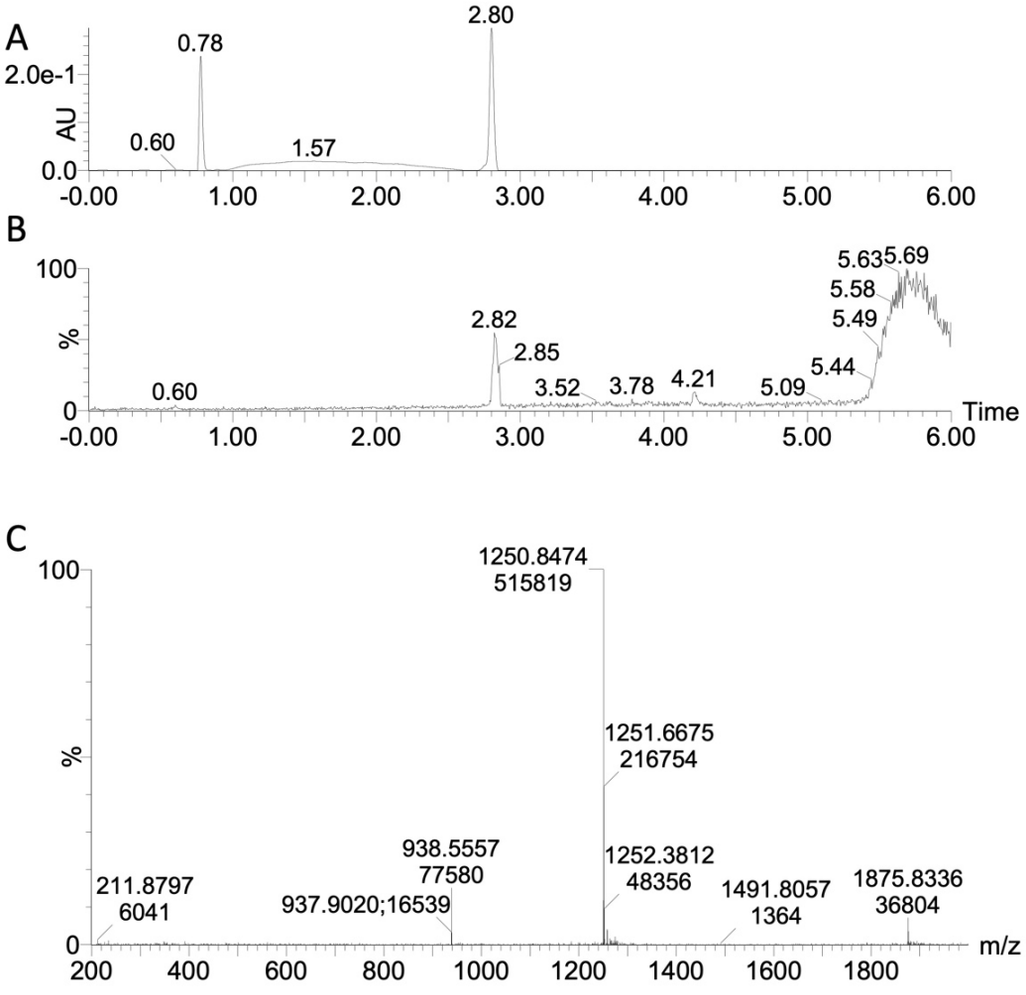


Figure 5.3. LC-MS of IPP5.

A) UV absorbance signal at 210 nm. The peak at 0.78 min is caused by the injection of the sample. It is due to the different percentage of acetonitrile in the sample in comparison to the LC gradient as well as the presence of salt in the sample. The x-axis is given in minutes. B) Total ion chromatogram (TIC). The x-axis is given in minutes. C) Mass spectrum of the peptide peak eluting at 2.82 min.

5.2 Solubility, Net Charge and pI are All Influenced by Lipidation

Initial experiments were used to measure the effect of lipidation on the solubility of the peptide. The solubility of IPP4 and IPP5 was determined as in the earlier studies on Am-GLP-1 and the results are shown in Figure 5.4A. The results establish that lipidation greatly reduces the solubility of the peptide, such that IPP4 and IPP5 are only soluble at pH values higher than their pIs, i.e., towards the basic pH range, Figure 5.4A. IPP5 is soluble from pH 6.5 upwards and IPP4 from pH 7.5. upwards. The net charge was calculated for both peptides in a similar way as for Am-GLP-1 and GLP-1 and as described in Section 2.15 and results are shown in Figure 5.4B. In addition, the net charge of Liraglutide was also calculated and is shown. Theoretically, under basic conditions, the net charge on IPP4 parallels the net charge on IPP5 with a net charge difference of plus one. The net charge of IPP4 is also similar to the net charge on GLP-1 and only shows differences at highly acidic or basic conditions.

The theoretical pI values for IPP4 and IPP5 were calculated using standard methods as described in Section 2.15 and are 5.5 for IPP4 and 4.7 for IPP5. The values were also determined experimentally and are 6.1 and 5.5 for IPP4 and IPP5, respectively, Figure 5.4C. The measured values are similar to, but not the same as, the theoretical values, suggesting that there is some structure or local charge effects in the lipidated peptides in solution which results in perturbed pK_a values for some side chains. The peak of IPP5 is broader and less focused than for IPP4, which could be caused by the formation of stable oligomers or the need for a longer focusing time. The latter was, however, restricted because of the pI markers running out of the detectable range of the capillary, Section 1.16.

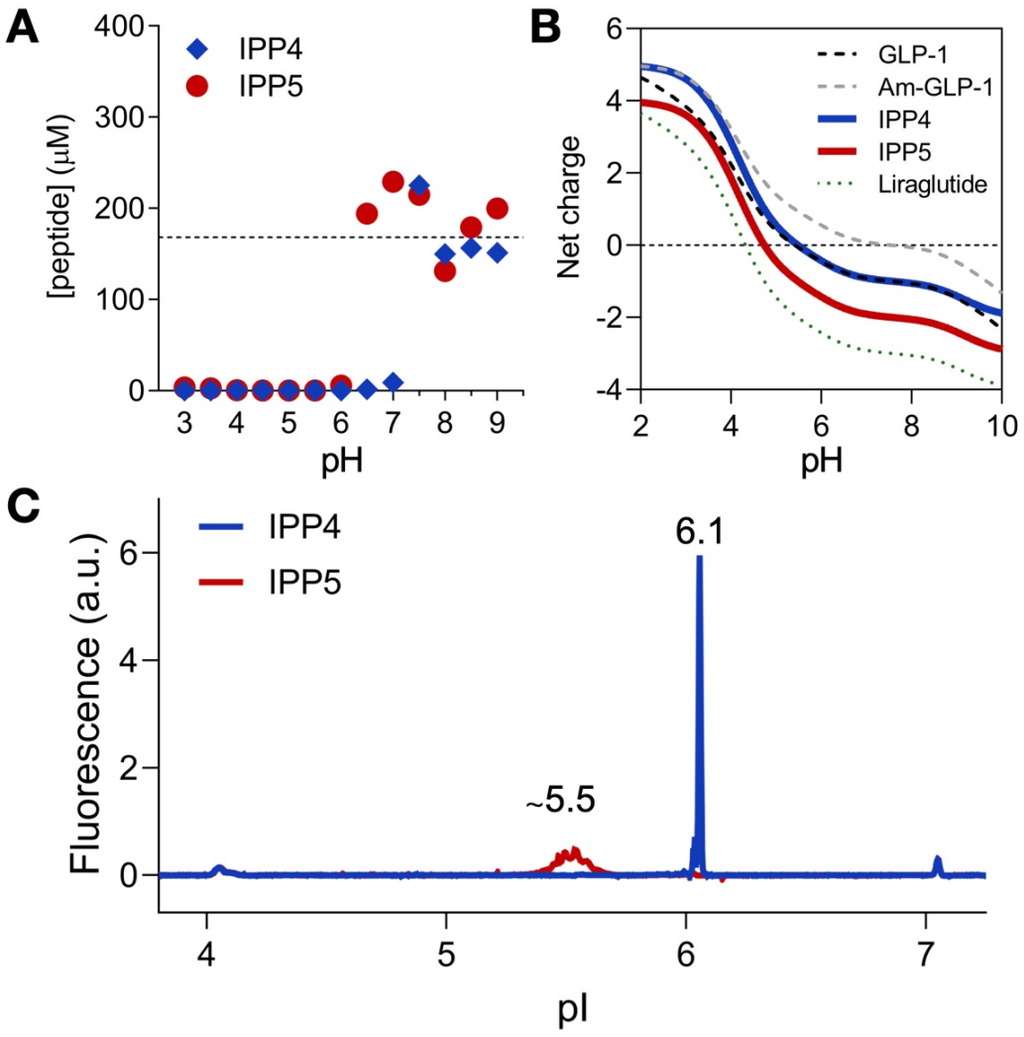


Figure 5.4. Solubility, net charge and experimentally determined p*I* of lipidated GLP-1 peptides.

A) Measured peptide concentration after dissolving approx. 0.25 mg of the peptide into 0.5 mL buffer at different pHs and after filtration of the sample with a 0.2 µm PVDF filter. IPP4 (red closed circles) and IPP5 (blue closed diamonds). B) Calculated net charge *versus* pH based on the peptide amino-acid sequence and standard pK_a values. The data for IPP4 (blue solid line) and IPP5 (red solid line) are shown in comparison to the calculated net charge of GLP-1 (black dashed line), Am-GLP-1 (grey dashed line) and Liraglutide (green dotted line). C) Experimental determination of the pI values for IPP4 (blue line) and IPP5 (red line). The pI was measured using capillary isoelectric focusing (iCIEF). The method used included urea to prevent precipitation.

5.3 A pH Induced Switch of Aggregation Propensity

For all pH values at which the peptides showed sufficient solubility, ThT aggregation assays were performed to learn more about the aggregation propensity of the two lipidated peptides at different pHs. IPP4 shows two regions which have different aggregation propensities, Figure 5.5A. At pH 7.5 and 7.7, the ThT signal shows a strong sigmoidal increase in fluorescence, whilst at pH 8.2 and 8.7, there is rather little increase in ThT fluorescence over the timeframe studied. Thus, if one assumes that the ThT fluorescence is linearly related to the formation of fibrils, IPP4 appears to have a much higher propensity to form fibrils at near neutral pH. IPP5 also shows a change in aggregation propensity with pH, Figure 5.5B.

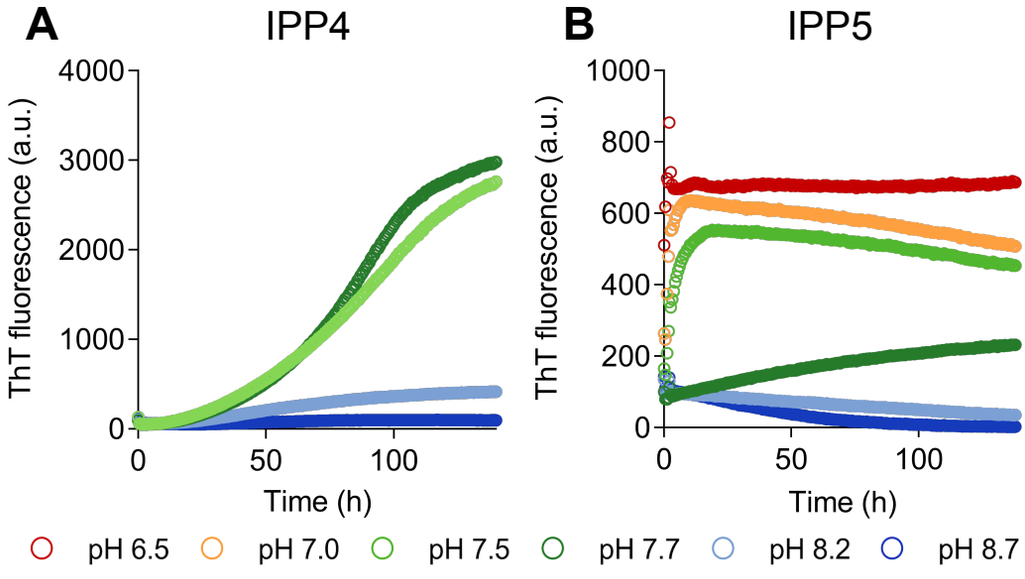


Figure 5.5. ThT aggregation assay data for lipidated peptides at different pH values.

Measurements of IPP4 and IPP5 conducted in 25 mM phosphate buffer at pH 6.5, (red open circles), pH 7.0 (orange open circles) and pH 7.5 (light green open circles) and 25 mM Tris buffer at pH 7.7 (dark green open circles), pH 8.2 (light blue open circles) and pH 8.7 (dark blue open circles). A) ThT assays of peptide aggregation for ~150 µM IPP4 from pH 7.5 to 8.7 at 37 °C. For each pH, the average of triplicate runs is shown. B) ThT fluorescence for ~150 µM IPP5 from pH 6.5 to 8.7 at 37 °C. For each pH, the average of triplicate runs is shown.

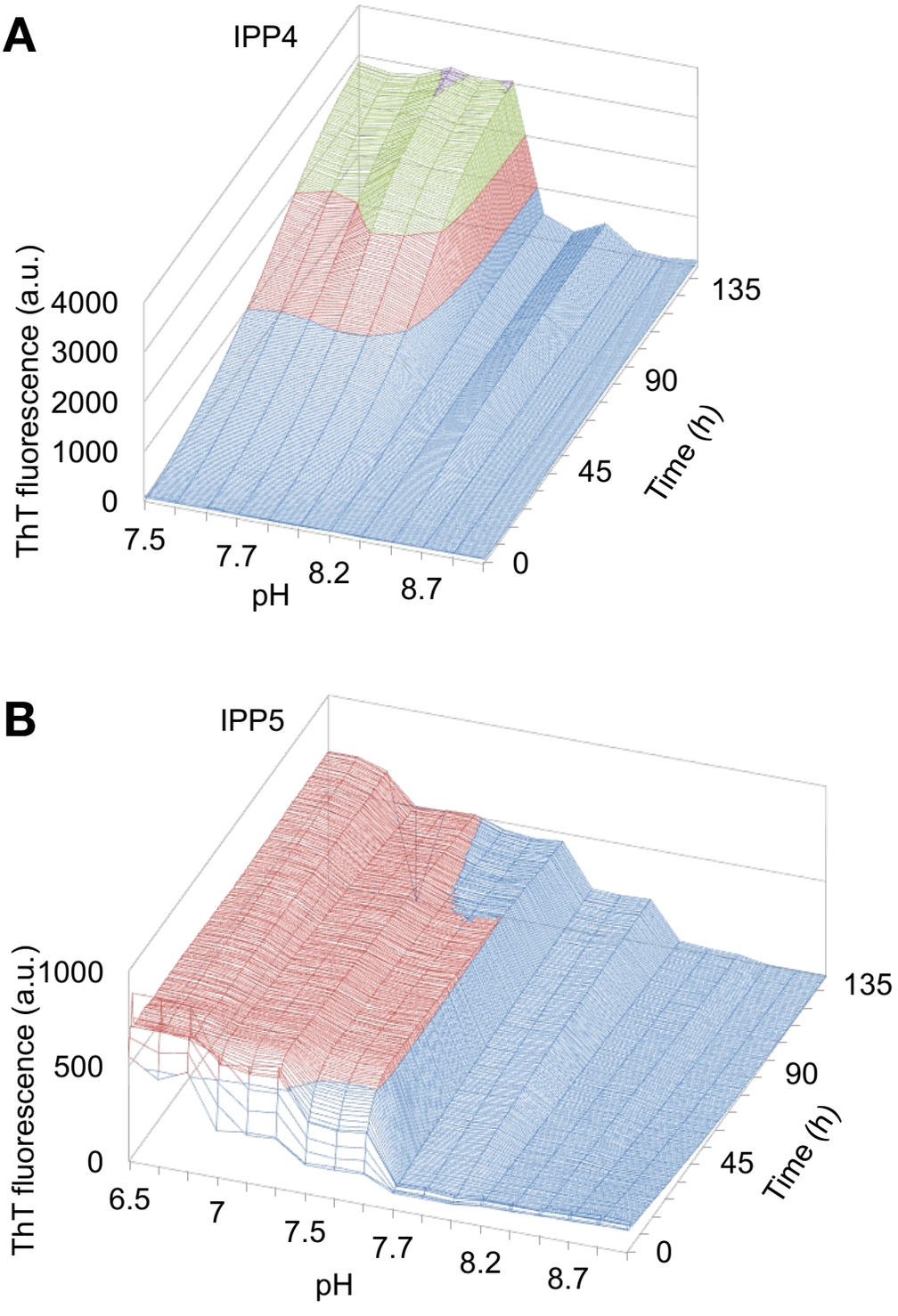


Figure 5.6. Three-dimensional aggregation landscapes illustrating the ThT assay data for peptide aggregation for lipidated peptides at different pH values.

A) ThT assays of peptide aggregation for 150 µM IPP4 from pH 7.5 to 8.7 at 37 °C. For each pH, the results of triplicate runs are shown. The x-axis shows the pH value, z-axis time (h) and the y-axis ThT fluorescence (arbitrary units) with 0-1000 (blue), 1000-2000 (red), 2000-3000 (green) and 3000-4000 (purple). B) ThT fluorescence for 150 µM IPP5 at pH 6.5 to 8.7 at 37 °C. For each pH, the results of triplicate runs are shown. The x-axis shows the pH value, z-axis time (h) and the y-axis ThT fluorescence (a.u.) with 0-500 (blue), 500-1000.

At pH values between 6.5 and 7.5, the samples show no lag phase and nearly no growth phase, but rapidly reach a moderately high ThT fluorescence plateau which may indicate fibril formation. However, on increasing the pH to 7.7, aggregation is significantly slower and by pH 8.2 and 8.7, there is relatively little fibril formation over the time course of the experiment as indicated by increased ThT fluorescence. Indeed, under these conditions, the ThT signal decreases not increases with time. It is worth mentioning that all samples of IPP5 from pH 7.5 to 8.7 show an initial ThT signal that does not start at zero, in contrast to the data for IPP4. Similarly, at pH 7.0 and 7.5, once the moderately high fluorescence signal has been reached, it then decreases slowly over time. 3D plots of the ThT data are shown in Figure 5.6A & B, illustrating the two different regions of aggregation behaviour for each peptide.

5.4 Aggregation Kinetics of Lipidated Peptides at Different Peptide Concentrations

The peptide concentration dependence of the aggregation of IPP4 and IPP5 at different pH values probed by ThT fluorescence measurements are shown in Figure 5.7. The ThT signal for IPP4 in phosphate buffer at pH 7.5 at higher peptide concentrations are sigmoidal in shape and data were fitted to Equation 2.4 and 2.5, to obtain values of $t_{1/2}$ and lag time. At lower peptide concentrations, 25 and 50 μM, the fits did not converge and therefore values for $t_{1/2}$ and lag time were not determined for these concentrations. Both kinetic parameters, $t_{1/2}$ and lag time, decrease with increasing peptide concentration, Figure 5.8A & C. IPP4 in Tris buffer at pH 8.2 also showed a sigmoidal increase in ThT signal; however, the amplitude of the change in signal was approximately four times lower than at pH 7.5, Figure 5.7B. These data were also fit to Equation 2.4 and 2.5, and values for $t_{1/2}$ and lag time shown in Figure 5.8B & D vary rather little with peptide concentration. It is interesting to note that there is a higher variance of the data at pH 8.2 compared with pH 7.5, Figure 5.7B.

In contrast to IPP4, the ThT kinetics of IPP5 aggregation at pH 6.5 starts at relatively high ThT fluorescence values but the signal does not change

significantly with time, Figure 5.7C. At pH 7.5, there is no lag phase, but a short sharp increase in ThT signal to some maximum is observed which is followed by a slow decrease in signal over longer times, Figure 5.7D. In both cases the maximum ThT signal for each peptide concentration is approximately linearly proportional to the peptide concentration, Figure 5.7C & D.

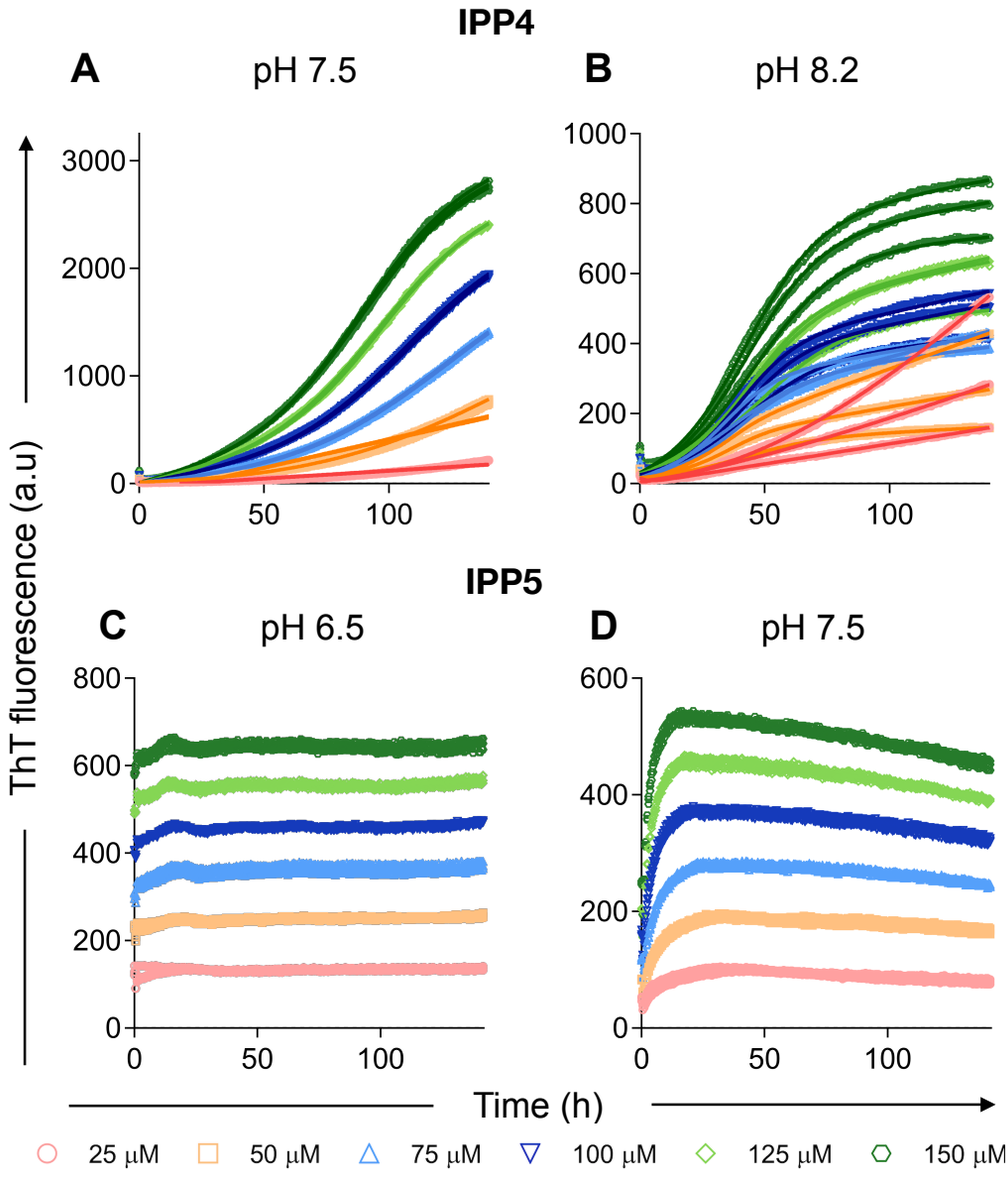


Figure 5.7. ThT aggregation assays of lipidated peptides at different peptide concentrations and pH values.

All samples were measured at the following peptide concentrations: 25 µM (pink), 50 µM (orange), 75 µM (blue), 100 µM (dark blue), 125 µM (green), and 150 µM (dark green). All experiments performed at 37 °C. A) IPP4 in 25 mM phosphate buffer at pH 7.5. The solid lines show the best fit of the data to Equation 2.4. B) IPP4 in 25 mM Tris buffer at pH 8.2. The solid lines show the best fit of the data to Equation 2.4. C) IPP5 in 25 mM phosphate buffer at pH 6.5. Data not fitted. D) IPP5 in 25 mM phosphate buffer at pH 7.5. Data not fitted.

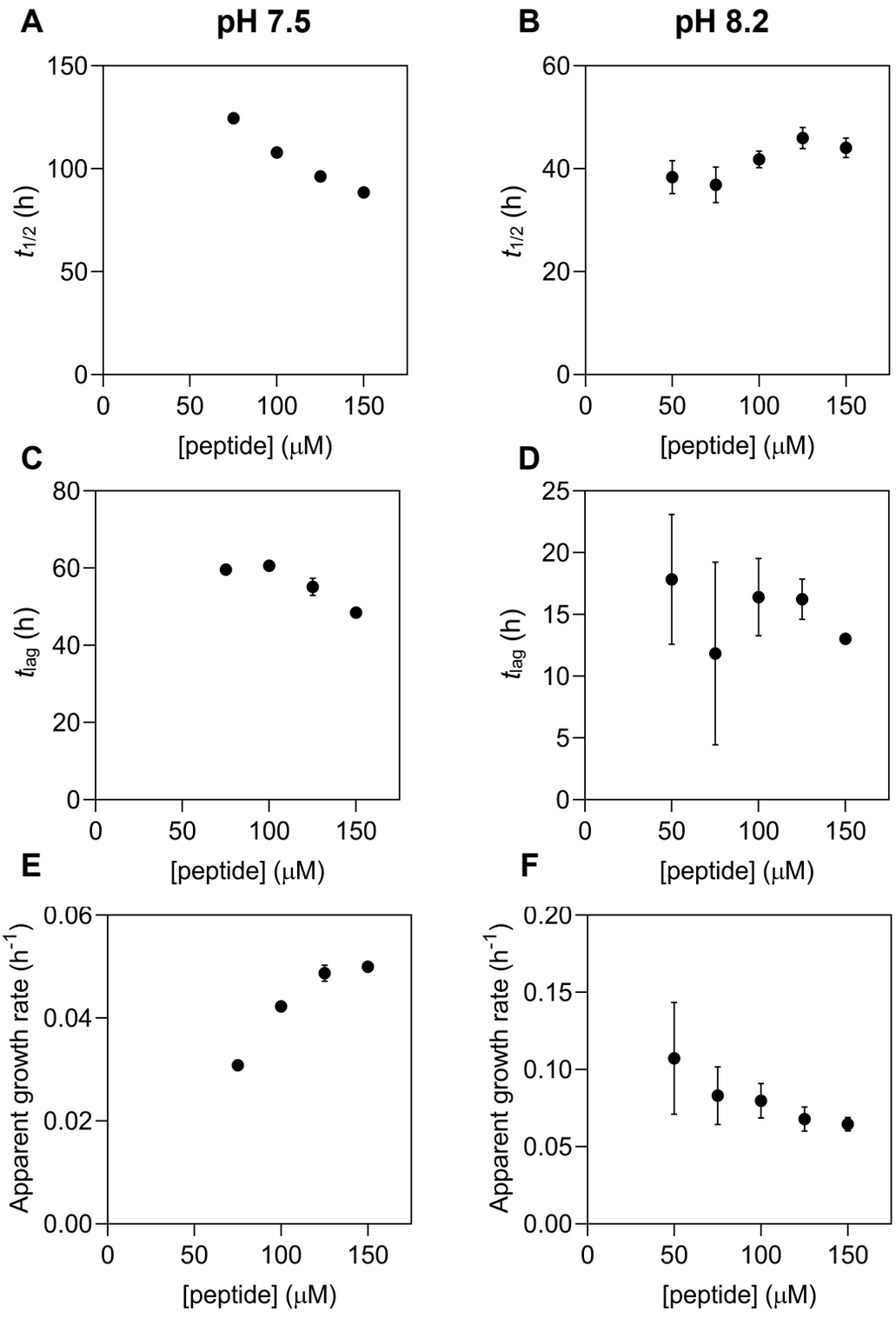


Figure 5.8. Kinetic parameters from the aggregation kinetics of IPP4.

ThT kinetic data for six different peptide concentrations between 25 – 150 μM were fitted to Equation 2.4. Fits at lower peptide concentrations did not converge and therefore values for $t_{1/2}$, lag time and apparent growth rate are not shown for 25 and 50 μM at pH 7.5 or 25 μM at pH 8.2. The error bars show the standard deviation between triplicate measurements. A) $t_{1/2}$ at pH 7.5. B) $t_{1/2}$ at pH 8.2. C) t_{lag} at pH 7.5. D) t_{lag} at pH 8.2. E) Apparent growth rate at pH 7.5. F) Apparent growth rate at pH 8.2.

5.5 Imaging of Higher-Order Structure in Aggregated Samples of Lipidated Peptides

The higher-order structures formed during the aggregation of the lipidated peptides, and proof of whether amyloid-like fibrils are forming or not, was assessed by SEM imaging. SEM imaging does not allow the highest resolution, but it gives a valuable overview of the higher-order structures of aggregates formed. Differences in the higher-order structure can point towards differences in the microstructure of the aggregates. Images of IPP4 aggregated in phosphate buffer at pH 7.5 recorded at different levels of magnification are shown in Figure 5.9. No fibrillar structures are observed at all at any level of magnification. Instead, roundish white dots on the barely covered region of the slide are seen, along with a sharp transition from the densely packed surface to the barely covered region, Figure 5.9C. The roundish white dots were analysed in detail: they were counted and the length of the major and minor axis of the dots were determined, Figure 5.10. It showed that the length of the major axis is Gaussian distributed with a mean of 14 nm, Figure 5.10C. Similarly, the length of the minor axis is distributed around 12 nm. Both values are linearly related to each other, Figure 5.10D.

IPP4 aggregated in Tris at pH 8.2 shows different higher-order structures in comparison to samples aggregated at pH 7.5, Figure 5.11. The images at low magnification show long (>100 μm) thin species, Figure 5.11A, however, these are unlikely to be fibrils as there are many breaks in them. In some cases, thicker species can divide into two species with a sharp and well-defined angle between them Figure 5.11A. At higher magnifications, the aggregates appear curved and even smaller diameter, ~20 nm, strands are observed, Figure 5.11B & C. In the right upper corner of Figure 5.11C a strand splits into two and the widths are approximately 20 nm or 10 nm for the thicker and the thinner strands, respectively. These structures show numerous cracks, probably caused by the drying process.

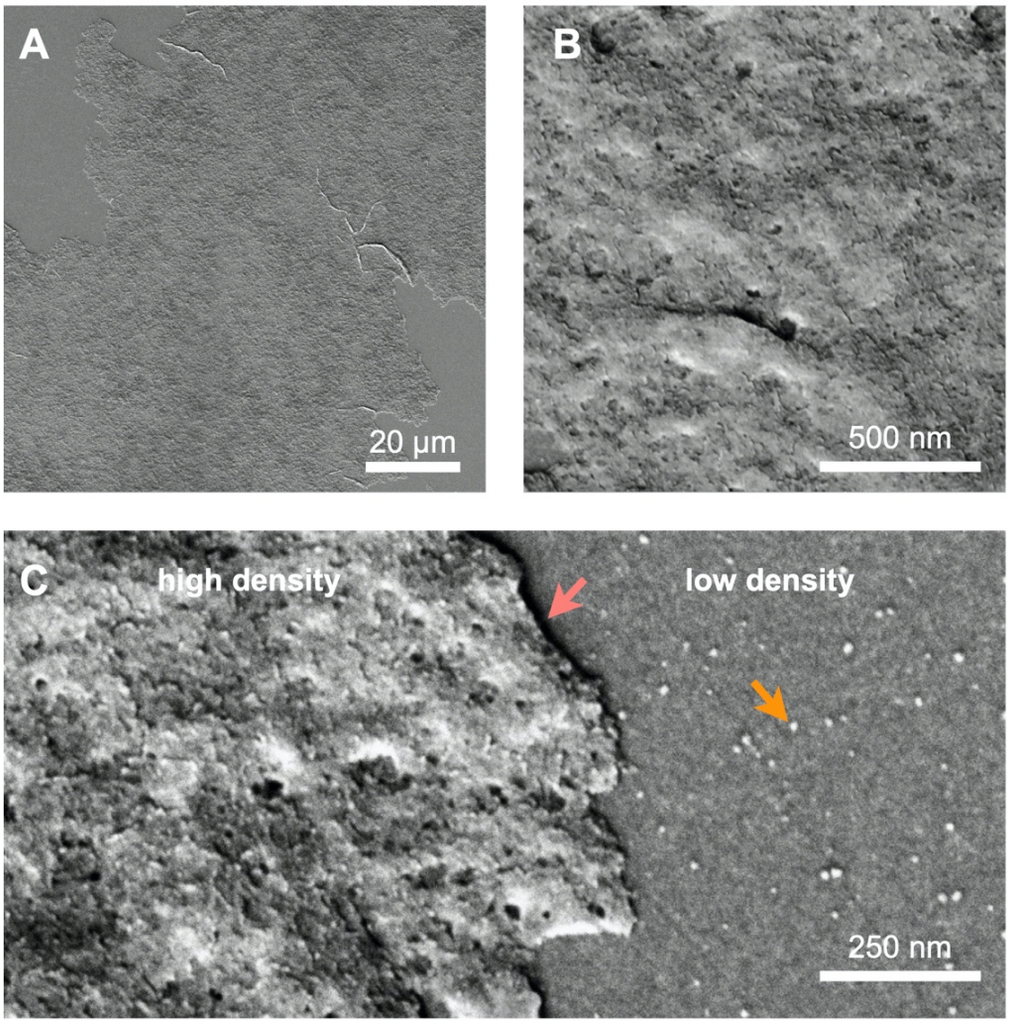


Figure 5.9. SEM images of aggregated forms of IPP4 in phosphate buffer at pH 7.5.

After 144 h of incubation time in the plate reader at 37 °C, 5 µL of 150 µM sample was dried on to a glass microscopy sheet. The images show the most typical of those obtained. A) Low magnification. Overview of regions with high and low peptide density. B) Higher magnification. Shows the image of a region where there is a high local concentration of aggregated peptide (Region 1). C) Highest magnification. High local density on the left and less densely populated regions of the surface on the right (Region 2). The border is marked with a red arrow. The less densely populated region shows white roundish species (orange arrow).

In regions with a high density of the aggregated sample, the surface is highly porous and no clear fibrillar structure can be observed, Figure 5.11D.

SEM images of samples of IPP5 before and after the aggregation process at pH 7.5 were taken, Figures 5.12 and 5.13. The sample imaged before the aggregation assay shows two different regions. One region is covered with more dense crystal-like aggregates, Figure 5.12A & C.

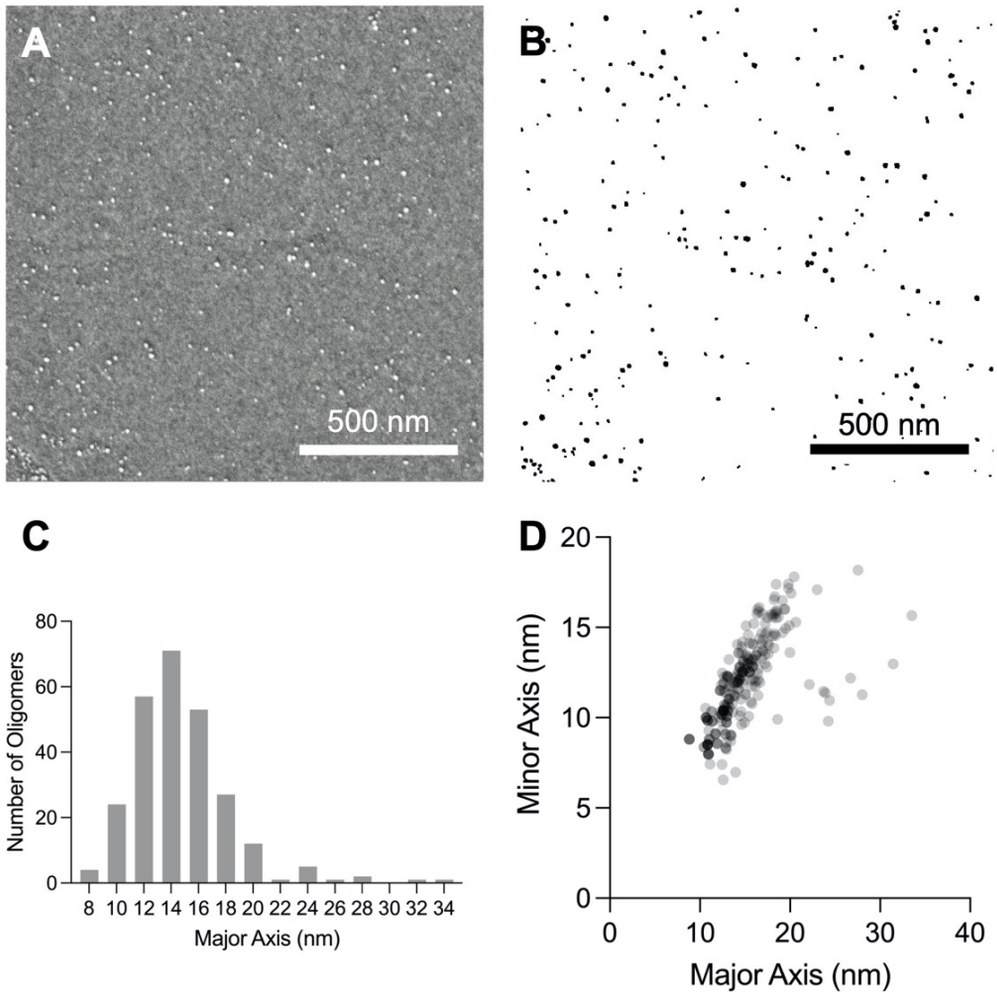


Figure 5.10. Size of oligomer-like IPP4 species formed in 25 mM phosphate buffer at pH 7.5

A) SEM image of less densely populated region with the white roundish species. B) Binary and inverted version of Image A. C) Histogram of the number of aggregates distributed by the length of the major axis of the aggregates. The data follows a Gaussian distribution. D) Relation between the length of the major and minor axis of the aggregates. The average length of the major axis value is approximately 14 nm and for the minor axis it is 12 nm.

From imaging of pure buffer these aggregates are most likely due to salt in the buffer, possibly mixed with sample but it is not possible to differentiate between them, Figures 4.14D and 4.15D. The second region shows a typical surface without dense peptide coverage, Figure 5.12B & D. The darker marks are most likely thin scratches on the glass surface. In general, the resolution of SEM is not high enough to say anything about the possible oligomeric structure of peptides before the aggregation assay.

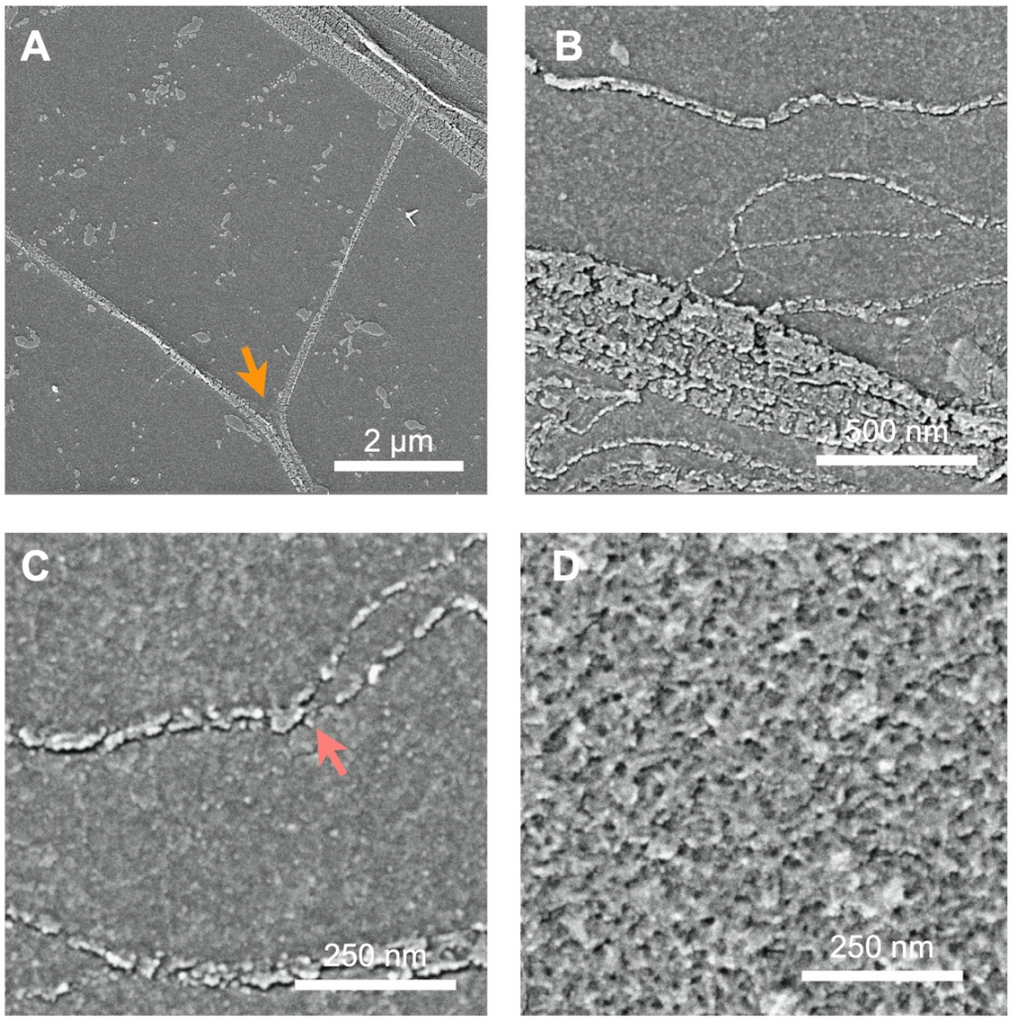


Figure 5.11. SEM images of aggregated forms of IPP4 in Tris buffer at pH 8.2.

After 144 h of incubation time in the plate reader at 37 °C, 5 µL of 150 µM sample was dried on a glass microscopy sheet. The images show the most typical of those obtained. A-C) Show images of regions where there is a low density of aggregates, but flow marks are visible. A) Sharp and well-defined angle between two splitting flow marks (orange arrow) C) A strand splitting into two (red arrow) D) Shows a densely populated region of the surface.

The same sample after the aggregation assay is shown in Figure 5.13. Two different regions were observed. A densely covered region shows a porous surface with some little holes in it, Figure 5.13A. The less densely covered region shows structures approximately 100 – 200 µm in diameter, seemingly consisting of shorter fibrils, as well as several thin short fibrils, Figure 5.13B & C. Characterisation of the fibrils showed a width of approximately 15 nm, Figure 5.13D. It is in the same range as the major axis of the roundish white dots or oligomer-like aggregates seen for IPP4 at pH 7.5, Figure 5.10.

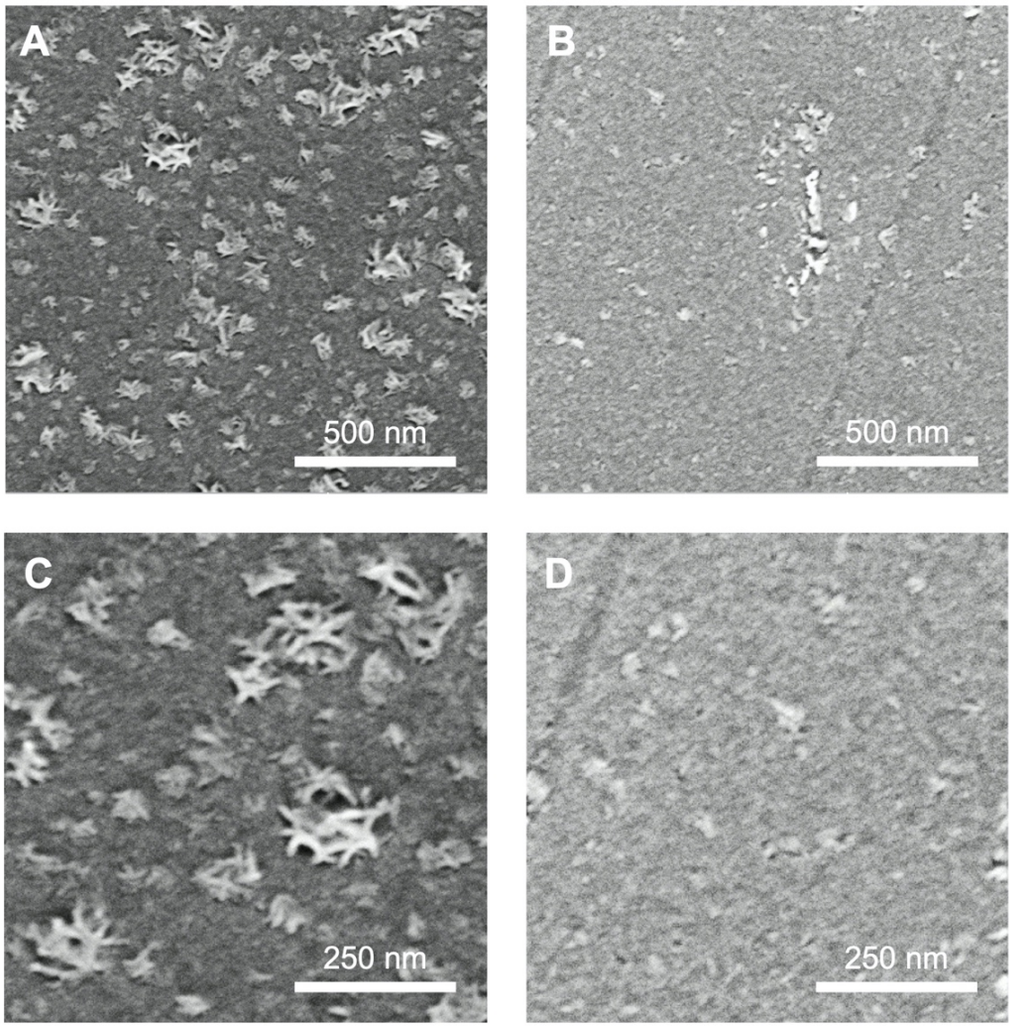


Figure 5.12. SEM images of freshly dissolved lipidated peptides of IPP5 in phosphate buffer at pH 7.5.

Before the samples were incubated in the plate reader, 5 µL of freshly prepared 150 µM IPP5 was dried on a glass microscopy sheet. The images show the most typical of those obtained. A & C) Show images of regions where there is a high density of aggregates. B & D) Shows a less densely populated region of the surface. C & D are at slightly higher magnifications than A & B.

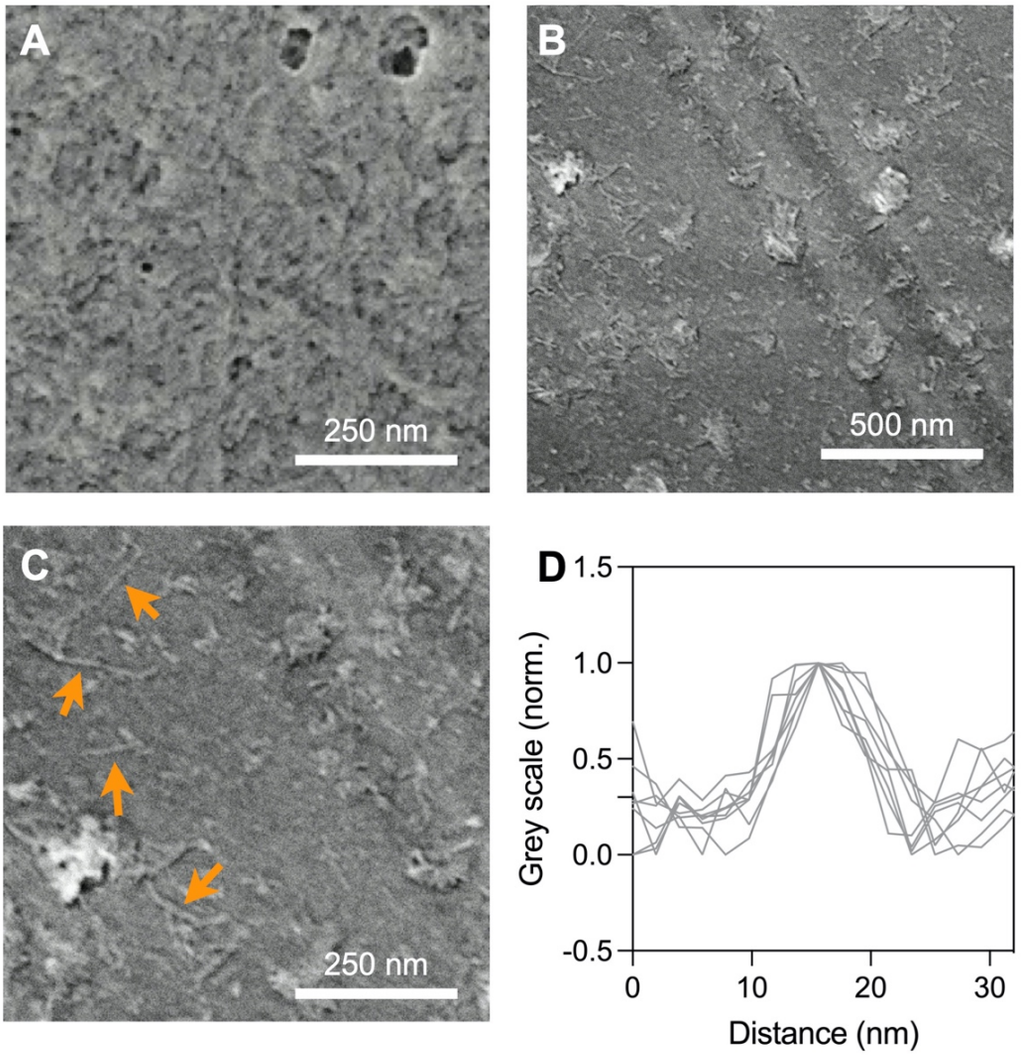


Figure 5.13. SEM images of aggregated forms of IPP5 in 25 mM phosphate buffer at pH 7.5 and measurement of width of fibrils.

After 144 h of incubation time in the plate reader at 37 °C, 5 µL of 150 µM sample was dried on a glass microscopy sheet. The images show the most typical of those obtained. A) Image of region where there is a high density of aggregates (Region 1). B) Less densely populated region of the surface (Region 2). C) Detail of Figure 5.13B. Arrows point towards measured fibrils. D) Normalised grey scale value plotted vs the distance orthogonal to the measured fibrils highlighted in Figure 5.13C. Each fibril was measured twice. The average width of the fibril is 15 ± 1 nm.

5.6 Secondary Structure of Lipidated Peptides

To learn more about the aggregates imaged with SEM a detailed study of the secondary structure was conducted. In Chapter 4, it was shown that even a small modification to GLP-1 (C-terminal amidation) influences the secondary structure of the peptide. It is, therefore, likely that a larger modification such as lipidation may also have an effect on secondary structure. The far-UV CD spectra of IPP4 and IPP5 were measured over a range of pH values in the same way as for GLP-1 and Am-GLP-1. However, the measurements on IPP4 and IPP5 were restricted to neutral and basic pH values due to the lack of solubility at lower pHs, Figure 5.4A.

The far-UV CD spectra of IPP4 were measured at pH 7.5 (phosphate buffer) and 8.0, 8.5 and 9.0 (pH of the Tris buffer at room temperature). At 37 °C, under aggregation conditions, the pH of the Tris buffers was 7.7, 8.2 and 8.7, respectively. Freshly prepared samples of IPP4 show only minimal differences in secondary structure between pH 7.5 and pH 9.0, Figure 5.14A. After the aggregation assay, the secondary structure of the samples aggregated at pH 8.2 and 8.7 show only minor changes in comparison to the freshly prepared samples, however, a clear change in the far-UV CD spectra occurs after aggregation at pH 7.5 and 7.7, Figure 5.14A. The CDPro analysis shows that this change in the far-UV CD spectra at pH 7.5 and pH 7.7 is due to the decrease of α-helical and increase of β-sheet structure. The data also show a small increase in turns, Figure 5.15A & B.

Before the aggregation assay, the far-UV CD spectra of IPP5 were measured from pH 6.5 to 9.0 at six different pHs, Figure 5.14B. The secondary structure directly after the peptide was dissolved in aqueous buffer shows that, under basic conditions, a similar α-helical content to IPP4 but decreasing pH from 8.0 to 6.5 results in a loss of helicity, Figure 5.15B. After the aggregation assay, the secondary structure, of samples aggregated between pH 7.7 and 8.7, changes only slightly in comparison to the secondary structure before the aggregation assay. However, between pH 6.5 and 7.5 there is a complete change in the secondary structure after aggregation. The sample at pH 7.7 shows intermediate behaviour, Figure 5.14B & D.

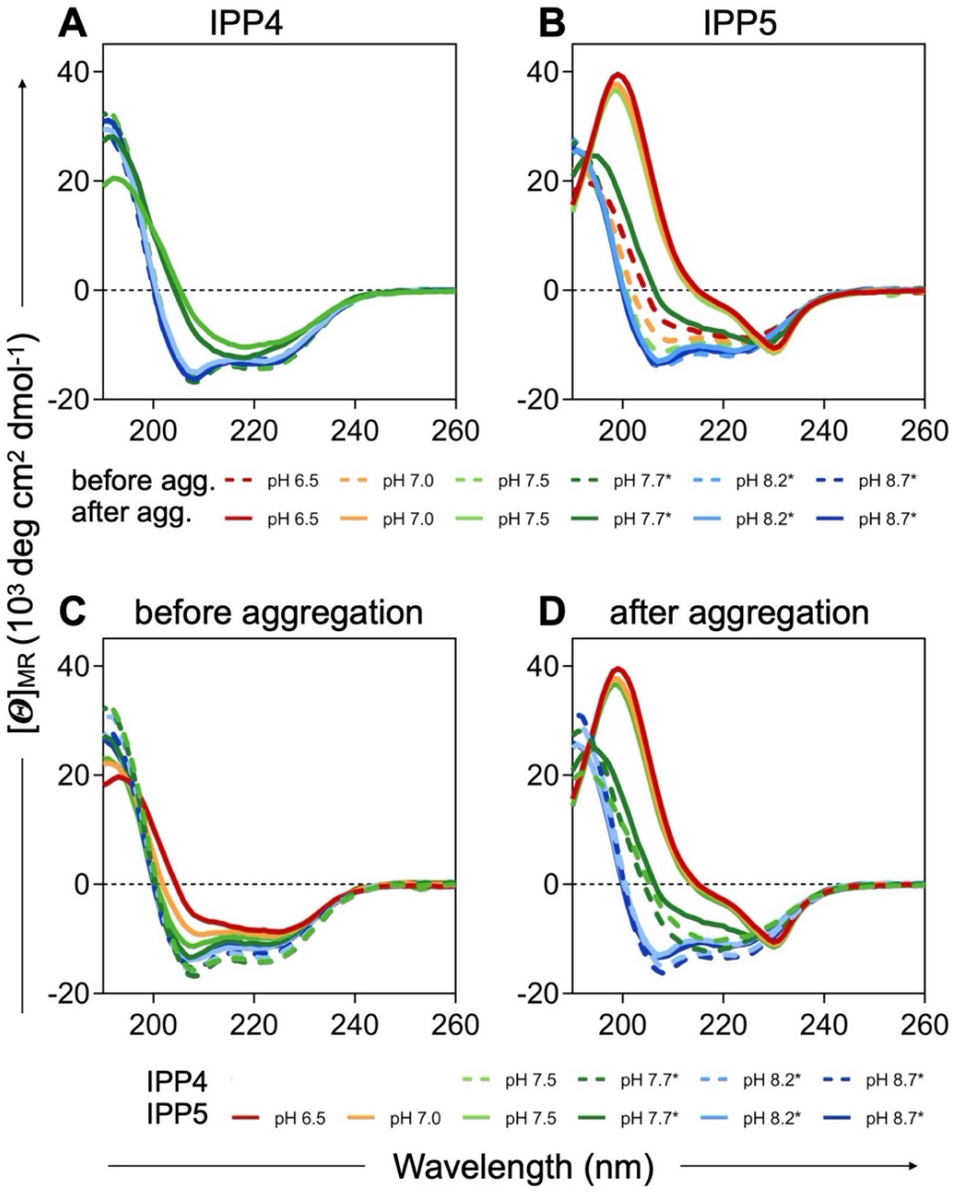


Figure 5.14. Far-UV CD spectra of lipidated peptides at different pH values before and after aggregation.

A) Far-UV CD spectra for IPP4 in 25 mM phosphate buffer at pH 7.5 (bright green) and 25 mM Tris buffer at pH 7.7 (dark green), 8.2 (bright blue) and 8.7 (dark blue) before (dashed lines) and after aggregation (solid lines). B) Far-UV CD spectra for IPP5 in 25 mM phosphate buffer at pH 6.5 (red), 7.0 (orange) and 7.5 (bright green) and Tris buffer at pH 7.7 (dark green), 8.2 (bright blue) and 8.7 (dark blue) before (dashed lines) and after the aggregation (solid lines). C) Comparison of the far-UV CD spectra of the same samples as shown in (A) and (B) before the aggregation assay. IPP4 data shown as dashed lines and IPP5 as solid lines. D) Far-UV CD spectra of the same samples as shown in (A) and (B) after 144 h aggregation in the plate reader under the same conditions as the ThT assay. IPP4 data shown as dashed lines and IPP5 as solid lines. *The far-UV CD spectra were recorded at 25 °C, while the aggregation assays were performed at 37 °C. Thus, the pH values in the Tris buffers were 0.3 pH units higher when the far-UV CD spectra were acquired compared to the pH at which aggregation took place in the plate reader.

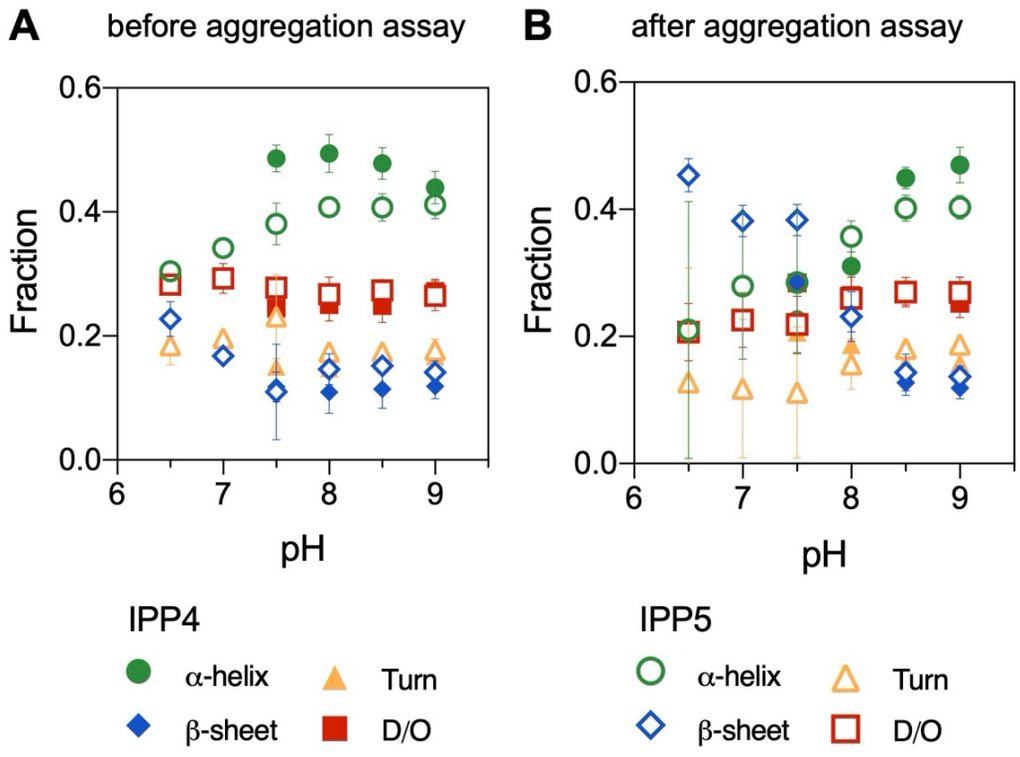


Figure 5.15. Secondary structure of the lipidated peptides at different pH values before and after the aggregation assay.

The fractions are calculated from the far-UV CD spectra shown in Figure 5.14 using CDPro. A) Fractions of α-helical, β-structure, turn and disordered structure of IPP4 and IPP5 after the peptides were dissolved, filtered and measured within 15 min. B) Fractions of α-helical, β-structure, turn and disordered structure of IPP4 and IPP5 after aggregation. The error is given as the standard deviation of the three algorithms used to calculate the secondary structure, Section 2.10. D/O stands for disordered structure.

If one compares the far-UV CD spectra before the aggregation assay of IPP4 and IPP5 directly, it can be seen that IPP5 has a less pronounced minimum at 208 and 220 nm and its maximum at 190 nm is lower compared to IPP4, Figure 5.14C. Plotting the far-UV CD spectra of the samples after the aggregation assay shows again the similarity in the behaviour of both peptides at basic pH values, Figure 5.14D. It is worth mentioning that the sample of IPP4 at pH 7.5 shows certain similarities with the spectra of IPP5 at pH 7.7.

The far-UV CD data were further analysed using CDPro. It shows that IPP4 immediately after it is dissolved has more α-helical structure at pH 7.5, 8.0 and 8.5 in comparison to IPP5. At pH 9.0, the amounts of α- *versus* β-structure are similar for both peptides. For freshly prepared IPP5, the fraction of α-helical structure is higher with increasing pH. After aggregation, the CDPro analysis

shows only minor changes of their α-helical content for both peptides incubated at pH 8.2 and pH 8.7. In contrast, from pH 7.7 towards more acidic pH values the fraction of α-helix decreases and β-structure increases for both peptides after 144 h of incubation in the plate reader. After the aggregation assay and at pH 7.5, IPP5 has a higher fraction of β-structure in comparison to IPP4, despite showing a lower ThT signal, Figures 5.5A & B and 5.15B.

5.6.1 Influence of Lipidation and Amidation on the Secondary Structure

A direct comparison of the secondary structure of all four peptides investigated is shown in Figure 5.16. Immediately after the peptides are dissolved Am-GLP-1, IPP4 and IPP5 show, under all conditions, more α-helical structure than β-structure. The fractional range of α-helicity spans from 0.3 to 0.5, whereas, the fraction of β-structure is always below 0.2, except for IPP5 at pH 6.5. GLP-1, the only non-amidated peptide of the four, shows a different behaviour of its secondary structure. At pH 6.5, the fractions of α-helical and β-structure for GLP-1 are still similar to the values of Am-GLP-1, IPP4 and IPP5. However, as the pH increases, the fraction of α-helices for GLP-1 becomes smaller and the fraction of β-structure higher, compared with the trend for the other three peptides where generally α-helicity increases with pH with a concomitant decrease in β. For GLP-1, from pH 8.0 upwards, the fraction of β-structure even exceeds the fraction of α-helical structure, Figure 5.16A.

The main observation after aggregation is that the α-helical fraction for both lipidated peptides stays at a high level at basic pH values, consistent with the SEM data which indicate that neither lipidated peptide form large amounts of fibrils under these conditions. This suggests that the lipidation stabilises the α-helical structure, especially at basic pH values. The non-lipidated peptides GLP-1 and Am-GLP-1 show α-helical fractions approaching zero and increased amounts of β-structure, consistent with the fact that these have formed amyloid-like fibrils over the 144 h incubation period.

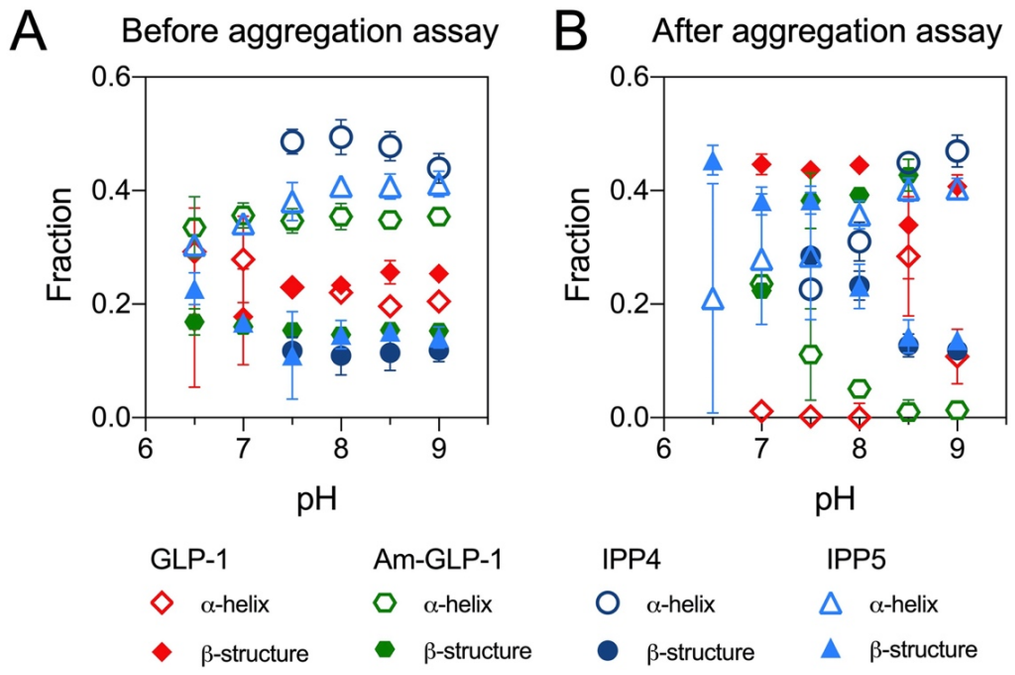
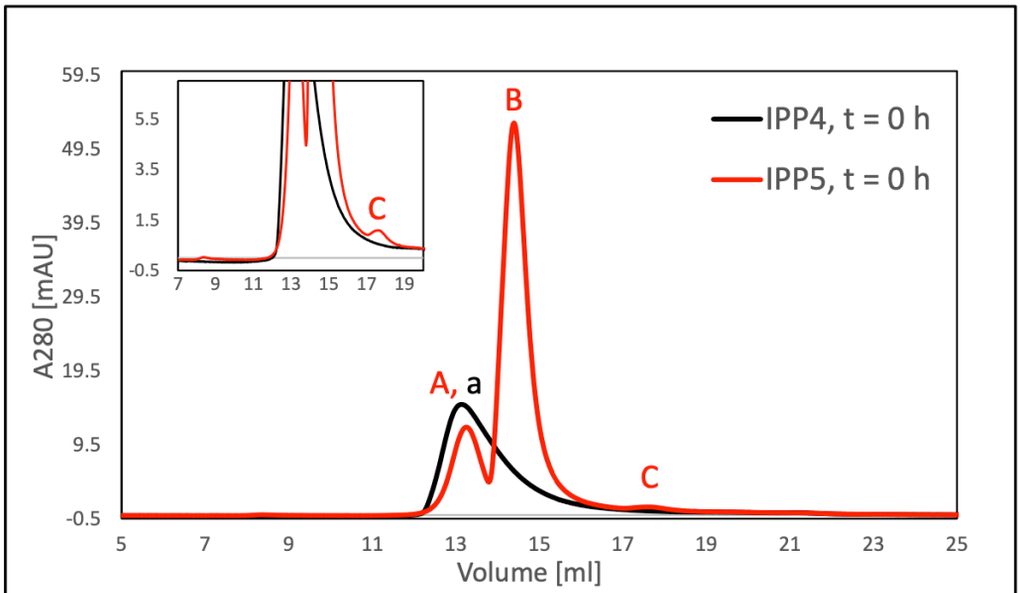


Figure 5.16. Secondary structure of GLP-1, Am-GLP-1 and the lipidated peptides at different pH values before and after the aggregation assay.

The fractions are calculated from the far-UV CD spectra shown in Figures 4.29, 4.30, 4.31 and 5.14 using CDPro. A) Fractions of α-helical and β-structure of GLP-1, Am-GLP-1, IPP4 and IPP5 after the peptides were dissolved, filtered and measured within 15 min. B) Fractions of α-helical, β-structure, turn and disordered structure of GLP-1, Am-GLP-1, IPP4 and IPP5 after aggregation. The error bars show the standard deviation of the three algorithms used to calculate the secondary structure, Section 2.10.

5.7 Size of Lipidated Peptide Oligomers

Size-exclusion chromatography (SEC) measurements were conducted by Eva Brichtová to analyse the oligomer distribution of the lipidated peptides.[341] SEC is a chromatographic method in which molecules are separated by their hydrodynamic radii with larger species eluting earlier from the column than smaller ones. IPP4 and IPP5 have been measured at a concentration of 135 µM in 25 mM phosphate, pH 7.5. Before each run, the samples were filtered to remove any larger insoluble aggregates. At the first time point, IPP4 and IPP5 show two different size distributions, Figure 5.17. IPP4 shows one broad peak eluting at 13.1 mL (A) (Figure 5.17), whereas IPP5 shows three different peaks eluting at 13.2 mL (A), 14.4 mL (B) and a minor peak at 17.6 mL (C), Figure 5.17.



IPP5	$V_{elution}$ [ml]	K_{av}	MW [Da]	n-mer (MW/$MW_{monomer}$)	Stoke's radius
A	13.2	0.324	88906	24	37.5
B	14.4	0.396	53191	14	31.3
C	17.6	0.596	12867	3	16.2

IPP4					
a	13.1	0.317	93380	25	38.1

$MW_{monomer}$	
IPP4	3777.79 Da
IPP5	3749.81 Da

Figure 5.17. IPP4 and IPP5 oligomer distribution analysed by SEC and estimated sizes of oligomers.

The SEC experiments were conducted by Eva Brichtová and are also described in her First Year Report. IPP4 and IPP5 samples at a concentration of 135 µM and in 25 mM phosphate buffer at pH 7.5 were analysed on a Superdex® 200 Increase 10/300 GL column. The samples were filtered prior to the SEC analysis. The molecular weight and Stoke's radius of oligomers were estimated using calibration curves.

The molecular weight and Stoke's radii were determined to get a better insight into the species present. However, the values obtained are rough estimates only because the commercially-available SEC calibration kits are purely based on globular proteins, Section 2.18. The estimated number of peptide subunits in the different oligomeric species are 24/25 for peak A, 14 for peak B and 3 for peak C. A comparison of these results with the literature shows that Liraglutide populates a similar-sized oligomer as IPP5. The studies of Wang *et al.* and Bothe *et al.* on Liraglutide, a non-amidated analogue of IPP5, identify

oligomers consisting of 12/13 peptide molecules at pH values below pH 7 which is similar in size to those found for IPP5 in peak B, however, under different pH conditions.[91,227] For Liraglutide, smaller oligomers comprising of approx. 7 monomers were found at more basic pH values.[91,227]

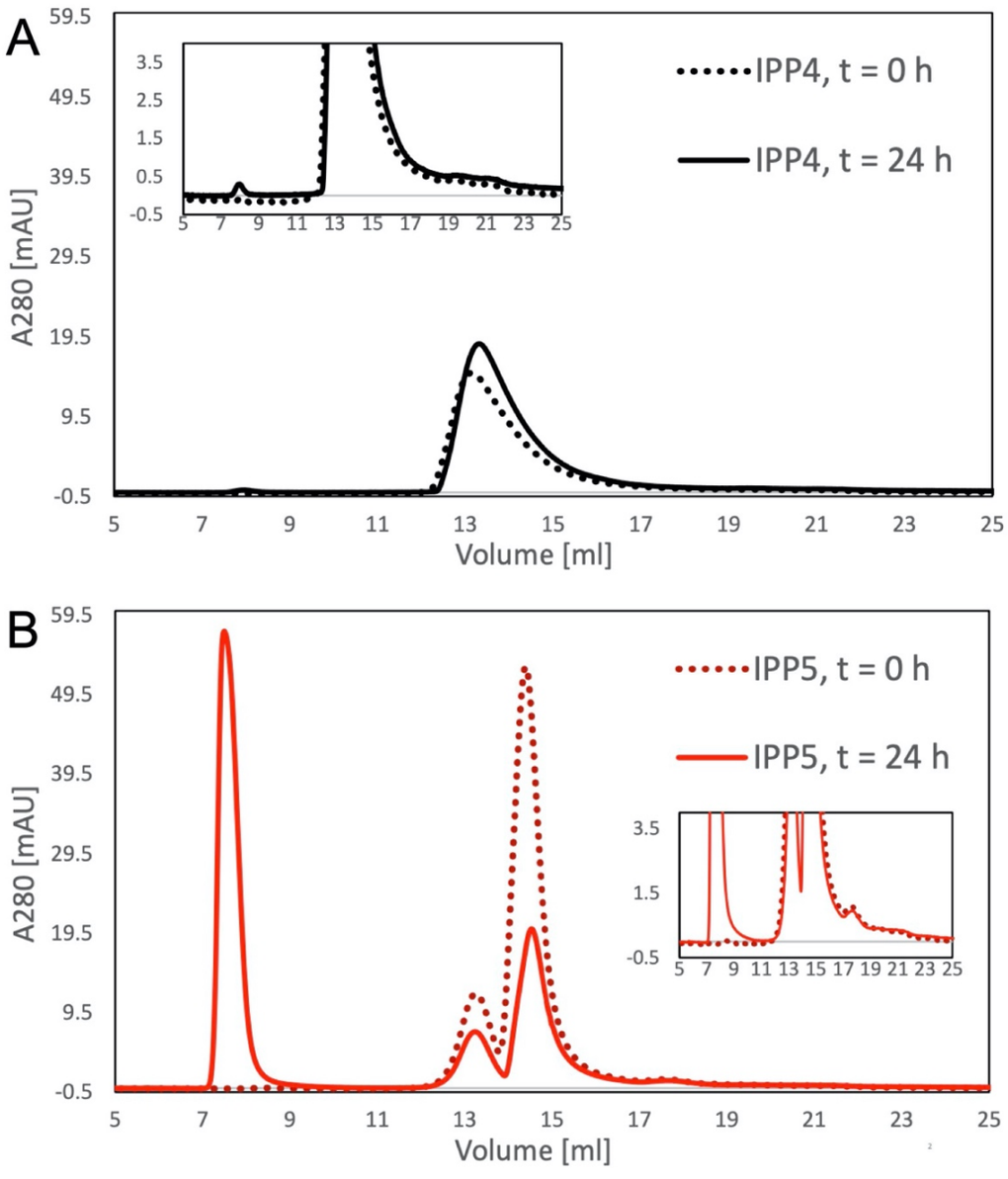


Figure 5.18. IPP4 and IPP5 oligomers distribution at two different time points analysed by SEC.

The SEC experiments were conducted by Eva Brichtová and are also described in her First Year Report. IPP4 and IPP5 samples at a concentration of 135 µM and in 25 mM phosphate buffer at pH 7.5 were analysed on a Superdex® 200 Increase 10/300 GL column. The samples were filtered prior to the SEC analysis. A) IPP4 measured at t = 0 h and after 24 h of incubation. B) IPP5 measured at t = 0 h and after 24 h of incubation.

Additional measurements were conducted on IPP4 and IPP5, to learn more about the influence of incubation on oligomer size and stability. After 24 h incubation in 25 mM phosphate buffer (pH 7.5) at 37 °C and agitation with 180 rpm, the oligomer distribution of IPP4 changed only marginally and only a minor shift of the maximum of the main peak was observed and a small new peak appeared, Figure 5.18A. This small peak elutes at around 8 mL and most likely results from the formation of higher molecular weight oligomers which elute in the void of the column, Figure 5.18A.

IPP5 analysed after 24 h of incubation shows an increase in a high molecular weight peak eluting at around 7.5 mL. Similar to IPP4 this peak is attributed to large M_w oligomers which elute in the void volume of the column. The results for IPP5 also showed a decrease in the oligomeric peaks eluting at 13.2 mL (A) and 14.4 mL (B), Figure 5.18B. The reduction of peak B is more pronounced than that of peak A, Figure 5.18B. It is possible that this oligomer (peak B) is on-pathway and crucial for the further self-assembly of IPP5 into larger species.

The faster self-assembly of larger soluble oligomeric species for IPP5 in comparison to IPP4 is in agreement with the ThT assay data which shows a steep increase in ThT fluorescence within the first 24 h only for IPP5 and not for IPP4 at pH 7.5, Figure 5.5A & B.

Dynamic light scattering (DLS) data were collected for 150 µM IPP5 in 25 mM phosphate buffer at pH 7.0 and 7.5 and in 25 mM Tris buffer at pH 8.5 and 9.0, Figure 5.19A-D. Note that the Tris buffer samples change their pH at 37 °C, under aggregation conditions, to pH 8.2 and 8.7. In all four samples, 60-80% of the intensity of the monomeric peptide is represented by the first peak (Peak I), Figure 5.20A. The actual percentage of the species represented by Peak I is expected to be even higher than the given percentage of the intensity as smaller particles scatter light less and have therefore a lower intensity in comparison to larger species. The apparent hydrodynamic diameters of these peaks are approximately 7 nm at pH 7.5, 8.5 and 9.0 (pH values at 25 °C) and 14 nm at pH 7, Figure 5.20B. The smaller hydrodynamic diameter suggests that the peptide at pH 7.5, 8.5 and 9.0 stays in the solution as monomers or assembles into small oligomers like dimers. The larger oligomers with a diameter of

approx.14 nm formed at pH 7 likely contain more monomeric units than those species formed under more basic conditions.

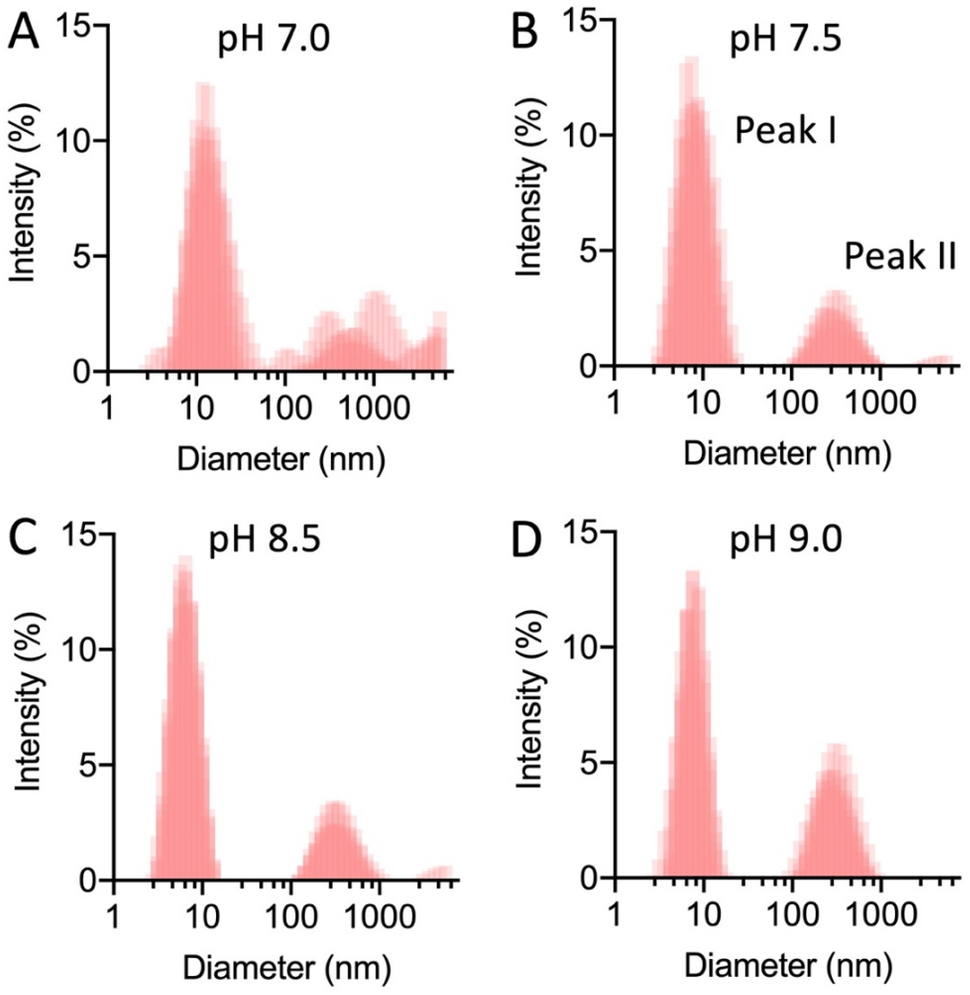


Figure 5.19. Results of DLS measurements to determine the size of IPP5 species in solution at time zero and at four different pH values.

DLS measurements of 150 µM IPP5 at 25 °C shown as intensity vs. hydrodynamic diameter. The darker pink areas show the overlap of three repeated measurements. A) 25 mM phosphate buffer at pH 7.0, B) 25 mM phosphate buffer at pH 7.5, C) 25 mM Tris buffer at pH 8.5 and D) 25 mM Tris buffer at pH 9.0. At 37 °C, under aggregation conditions, the Tris buffer pH values change to pH 8.2 and 8.7.

At all pH values studied, a smaller fraction of the total amount of peptide also forms oligomeric species with approx. hydrodynamic radii between 200 and 300 nm (Peak II). Here, a difference can again be seen between pH 7 and the IPP5 at more basic pH values, where the former shows a higher size variance of the larger aggregates.

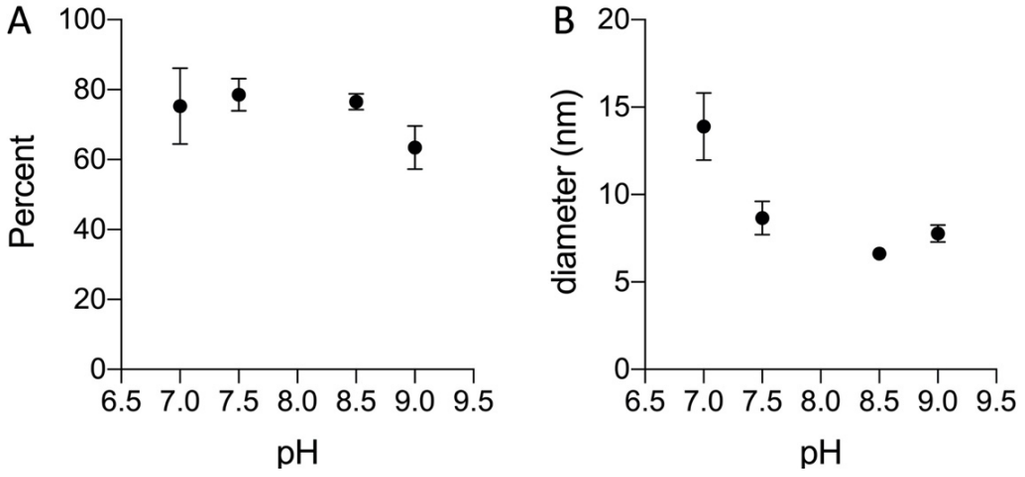


Figure 5.20. Analysis of DLS results: size and amount of smaller species of IPP5 formed at four different pH values.

Percentage of intensity and diameter of aggregates of peak I observed by DLS measurements of samples of 150 μM IPP5 at 25 °C. IPP5 was measured in 25 mM phosphate buffer at pH 7.0, 25 mM phosphate buffer at pH 7.5, 25 mM Tris buffer at pH 8.2 (37 °C) and 25 mM Tris buffer at pH 8.7 (37 °C). A) Percentage of IPP5 found in peak I at different pH values. B) Apparent hydrodynamic diameter of species in peak I vs pH.

After 144 h of incubation of IPP5 in a ThT plate reader at pH 7.0 and 7.5, the main peak observed in the DLS moves and indicates an increased amount of species with hydrodynamic diameters of around 200 nm, Figure 5.21A & B. These aggregates are likely either short fibrils or oligomers which have β-structure because at pH 7.0 and 7.5 no distinct lag time was observed in the ThT kinetics, Figure 5.5B, and similarly sized aggregates are already present at the first time point. Under these conditions, the signal corresponding to the smaller species with hydrodynamic diameters of approximately 7 or 14 nm, respectively, has nearly entirely vanished. At pH 8.2 and 8.7, the DLS measurements were not reproducible and, for unknown reasons, showed single spikes with high variance, Figure 5.21C & D. This may be because of the concentration of peptide used in these experiments, which was chosen to be consistent with other biophysical measurements, but which is rather low for DLS.

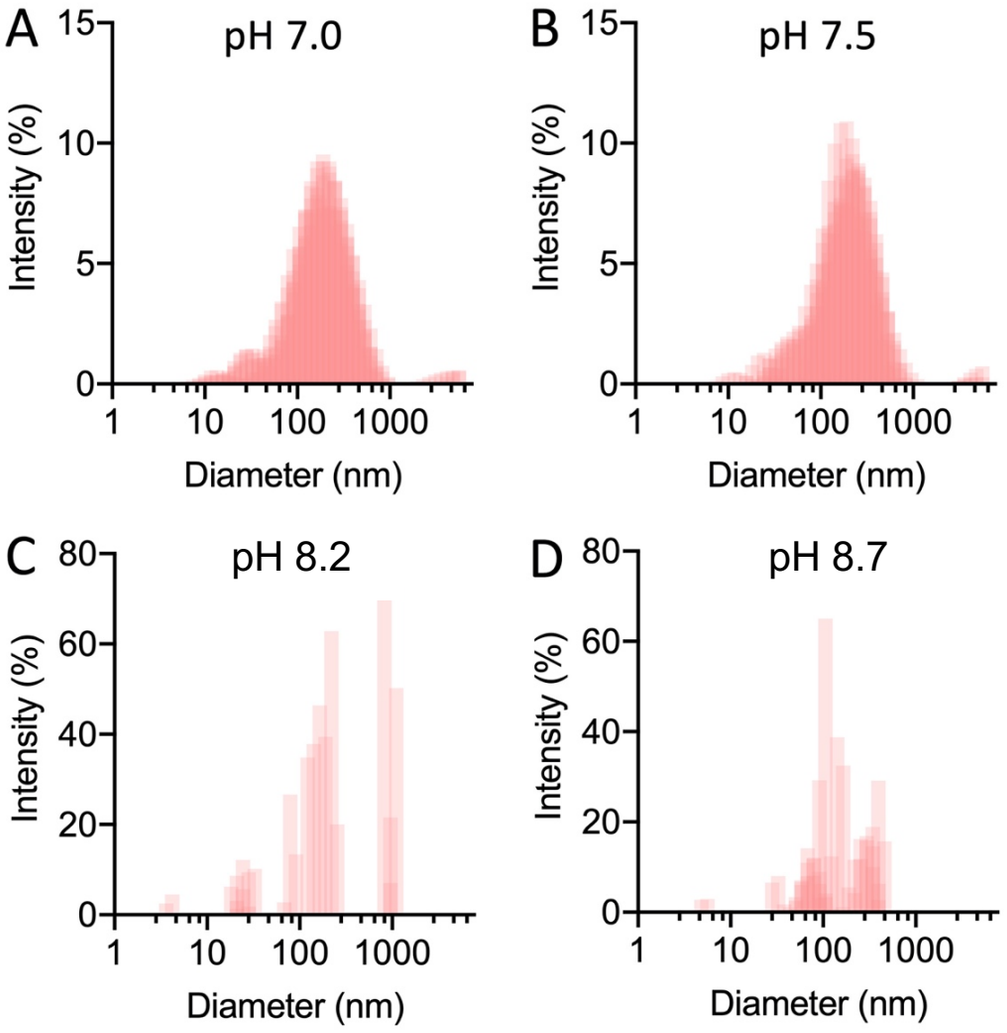


Figure 5.21. Results of DLS measurements to determine the size of IPP5 species in solution after 144 h incubation at 37 °C and at four different pH values.

DLS measurements performed at 25 °C of samples of 150 μM IPP5 which have been incubated with shaking at 37 °C. Data shown as intensity vs. hydrodynamic diameter. The darker pink areas show the overlap of three repeated measurements. A) 25 mM phosphate buffer at pH 7.0, B) 25 mM phosphate buffer at pH 7.5, C) IPP5 incubated in 25 mM Tris buffer at pH 8.2 and D) IPP5 incubated in 25 mM Tris buffer at pH 8.7.

5.8 Discussion

5.8.1 Solubility of Lipidated Peptides is Restricted to Neutral and Basic Conditions

The solubility of IPP4 and IPP5 is dominated by the influence of the lipid, which causes very low solubility at acidic pHs. Towards neutral and more basic pH values, the solubility increases for both peptides. For IPP4 this starts around pH 7.5, in comparison to pH 6.5 for IPP5. This difference is probably due to the differences in net charge at these pH values, where IPP4 has a smaller net charge than IPP5. The measured pI values of pH 6.1 and 5.6 for IPP4 and IPP5, respectively, are in good agreement with the solubility data, a significant increase in solubility being observed 1 to 1.5 pH units above the pI. The slightly higher difference of 1.4 pH units for IPP4 can be explained by the different slope of the net charge *versus* pH. The relative net charge of IPP5 increases faster than for IPP4 with increasing pH in this range.

5.8.2 Is There Evidence that Lipidated GLP-1 Peptides Form Amyloid-Like Fibrils?

IPP4

For IPP4, the fact that there is a sigmoidal increase in ThT fluorescence over time when it is incubated at 37 ℃ at pH 7.0 and 7.5 and the fact that the final ThT fluorescence is higher than that of GLP-1 and Am-GLP-1 under similar conditions, indicate that IPP4 might form fibrils under these conditions. In support of this, the far-UV CD spectra show clear increases in β-structure at pH 7.0 and 7.5. However, the SEM images recorded at the end of the incubation at pH 7.5 do not show any evidence of fibril formation. Nevertheless, there are regions which show dense peptide coverage possibly due to amorphous aggregation, indicating that some aggregation phenomenon has taken place. The density of coverage and the amorphous aggregation may well obscure any fibrils present in the sample.

In contrast, IPP4 at pH 8.2 after incubation at 37 ℃ for 144 h, shows a considerably lower ThT fluorescence either indicative of the formation of a smaller quantity of amyloid-like fibrils or a different species of aggregate that cannot bind ThT, or can bind only weakly. These results are consistent with the observed lack of change in the far-UV CD spectra of IPP4 before and after aggregation suggesting that amyloid-like fibrils are not formed at this pH. The SEM results are consistent with this: no amyloid-like fibrils are seen, however, there is some evidence of aggregated species of unknown type.

IPP5

For IPP5, an increase in ThT fluorescence was observed after incubation at 37 ℃ for 144 h at pH 6.5 and 7.5, albeit to a much smaller degree than IPP4 under similar conditions. The lower ThT fluorescence is despite the fact that IPP5 has a higher negative net charge in comparison to IPP4 under basic conditions which should attract the positively charged ThT to a greater extent. It has been shown that the ζ-potential of fibrils, which is strongly correlated to the net charge, influences the binding affinity of the positively charged Thioflavin T.[291] Despite the smaller change in ThT fluorescence, there was a noticeable change in the far-UV CD spectra of IPP5 under these conditions, the CDPro analysis indicating a pronounced increase of β-structure and a decrease in α-helical structure. At pH 7.5, the aggregated form of IPP5 showed some evidence of fibrils in the SEM imaging, however, the observed fibril-like structures differ from the typical rigid fibrils observed for GLP-1 and Am-GLP-1, Figure 5.13C. In this case, the IPP5 aggregates resemble more closely the thin curly fibrils observed in the early stages of aggregation of GLP-1, Figure 3.15B.

At pH 7.7 and above, there is little change in ThT fluorescence from start to end of aggregation, and the difference in the far-UV CD spectra before and after aggregation is not as pronounced as for the more acidic pH values. Thus, there is little evidence that amyloid-like fibrils have formed under these conditions.

5.8.3 ThT Fluorescence Caused by Non-Fibrillar Structures

Although an increase in ThT fluorescence is frequently used as an indicator of amyloid-fibril formation, it cannot be excluded that other aggregates, such as oligomers, can also interact with and increase the fluorescence of ThT. This is because ThT does not recognize fibrils *per se*, but it is assumed to bind to the molecular grooves which are typically found on the surface of amyloid fibrils.[362] This raises the question of the whether the increase in ThT fluorescence observed for both IPP4 and IPP5 at near neutral pH values is due to fibril formation or the formation of some other aggregated species.

The question arises, therefore, what is causing the high fluorescence signal if not amyloid-like fibrils? From the SEM imaging, it appears that IPP4 at pH 7.5 shows two different kinds of aggregated morphologies. The glass sheet was mostly covered by a dense rather amorphous cover of peptides, which constituted most of the peptide molecules. However, small oligomers could be observed in the less densely covered regions, Figures 5.9C and 5.10. It is possible that the densely covered region consists out of tightly packed oligomers and it is, therefore, not clear if there are two different species present or not.

At pH 8.2 the SEM images again show an extended region with thick peptide cover and in a less densely covered region, some extended structures can be seen, however, they do not show typical rigid fibril-like structures either. In both cases, ThT may bind to the non-amyloid like aggregates observed (amorphous or with more regular structure) and increase in fluorescence giving rise to the typical sigmoidal ThT kinetics observed. The same may also be the case for IPP5, however, here some short thin fibrils were also observed. The hypothesis that ThT also binds to non-fibrillar species, also explains why the maximum ThT fluorescence of IPP5 is lower than for IPP4 despite it forming more β-structure during the aggregation process.

Collectively, the results for both lipidated peptides suggest that during aggregation the peptides increase their β-structure but that this structure forms inside oligomers or other types of amorphous structures instead of amyloid fibrils.

5.8.4 Kinetics of the Formation of Aggregates

The ThT data for IPP4 and IPP5 at different pH values establishes that there are three types of behaviour observed. The first, in which the ThT fluorescence increases to a high value showing standard sigmoidal kinetics, is observed for IPP4 at pH 7.5 and 7.7. In the second class, the ThT shows an immediate or rapid increase in fluorescence reaching a final value that is significantly lower than in the first case. IPP5 shows such kinetics at pH 6.5 to 7.5. In the third case, there is either little or no increase in ThT fluorescence over the time course of the experiment, and in some cases, the low starting ThT fluorescence decreases even further. This is found for IPP4 and IPP5 at pH 8.2 and 8.7. The first behaviour observed suggests an aggregation reaction that follows a classic nucleation-propagation mechanism, the second behaviour is also consistent with this, but the initial primary nucleation must be rapid. However, in the second case, the kinetics are also compatible with a mechanism in which oligomerisation does not involve a nucleation step. The third case suggests that no aggregation occurs or that the aggregated species is not capable of binding to ThT.

5.8.5 Unusual Decrease of ThT Fluorescence

At pH 7.5 and higher peptide concentrations, the ThT fluorescence of IPP5 decreases linearly after reaching a maximum early in the assay, Figure 5.7D. It is interesting to note that pH 7.5 is close to the pH value where a change in aggregation behaviour is observed (pH 7.7 and higher). It is possible to speculate that if two different aggregated species with either different ThT affinities or which affect the ThT fluorescence differentially are populated under these conditions, that this slow decrease in the ThT signal, could be due to the transition from one aggregated species to another, Figure 5.22. Alternatively, the decrease could be explained by a structural change of the aggregate and therefore, a different ThT affinity or effect on fluorescence, Figure 5.22. From the amorphous structures observed with SEM for IPP5 at pH 7.5, it is likely that these are all non-fibrillar aggregates, Figure 5.13A.

It may also be possible that hydroxylation of ThT causes the decrease of ThT fluorescence observed at the higher pH values.[301]

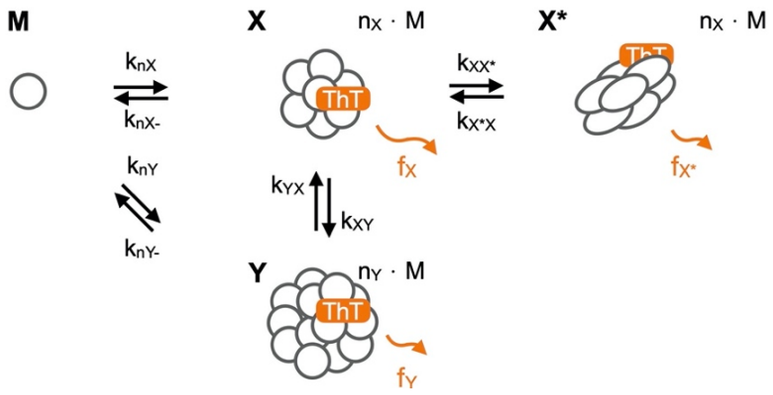


Figure 5.22. ThT fluorescence caused by non-fibrillar structures.

The formation of oligomers X and Y is described by the primary nucleation rate constants k_{nX} and k_{nY} as well as the dissociation rate constants k_{nX-} and k_{nY-}. The transition between oligomers X and Y and the structure change between oligomer X and X* is defined by the rate constants k_{XY}, k_{YX}, k_{XX*} and k_{X*X}, respectively. If smaller non-fibrillar aggregates bind ThT, as suggested by the IPP4 and IPP5 data, a change from one species X to another species Y can cause a change in ThT fluorescence f. Similarly, a conformational change illustrated by the change in structure from X to X* can have the same effect.

5.8.6 Peptide Concentration Dependence of the Aggregation Kinetics

IPP4

The aggregation kinetics at different peptide concentrations of IPP4 showed two distinct characteristics. At pH 7.5, the peptide aggregated faster with increasing peptide concentration, consistent with a classic nucleation-propagation mechanism. Under more basic conditions (pH 8.2), the kinetics change and slightly shorter but constant lag times show that the aggregation process must be influenced by an additional factor next to the nucleation-propagation mechanism, for example, off-pathway species. The observed decrease of the ThT intensity under basic conditions is most likely influenced by a change of the populated species, a structural change of the aggregate or the hydroxylation of ThT[301] as discussed in Section 5.8.5.

IPP5

The concentration dependence of IPP5 aggregation was measured at pH 6.5 and 7.5. In both cases, no lag time is visible in the aggregation kinetics which may be due to the lag phase being shorter than the time needed to prepare the samples before the aggregation assay is started or may be due to seeds in the freshly prepared samples. There is an approximately linear relationship between the final ThT signal and the peptide concentration, suggesting some type of aggregation process has occurred. It is most likely not caused by fibrils but ThT binding to oligomers, assuming that oligomers form rapidly after the peptide is dissolved in aqueous buffer and can bind to and restrict the rotation of the ThT leading to an increase in fluorescence.

5.8.7 Different Oligomer Sizes Influenced by Sequence and pH

The SEC measurements of IPP4 and IPP5 at pH 7.5 established that the two peptides can form oligomeric species, A, B and C. IPP4 forms only oligomers of type A which contain approx. 24 monomers. Oligomers of the same size were also found in smaller amounts for IPP5, the majority of which forms smaller oligomers containing approx. either 14 or 3 monomers.

The different-sized oligomers of IPP5 can be compared to oligomers found for Liraglutide, a non-amidated analogue of IPP5. Frederiksen *et al.* showed by small-angle X-ray scattering (SAXS) that at pH 8.1, Liraglutide assembles into oligomers with a global shape of hollow elliptical cylinders which are either hexa-, hepta- or octamers depending upon peptide concentration.[363] Asymmetric flow field flow fractionation-UV-multi-angle light scattering (AF4-UV-MALS) measurements confirmed the size of the oligomers as hexamers – the solutions being at lower concentrations in this case.[363] Additional simulations predicted heptamers as the most likely size of the oligomers.[363]

These findings were reproduced by Wang *et al.,* who also described one additional larger oligomer using static-light scattering (SLS).[227] They described a pH-dependent assembly of octamers below pH 7.0 and dodecamers under more basic conditions. Measurements were performed in a range between pH 6.4 and 8.7 at a concentration well over that used in the previous study (over 1

mM). Bothe *et al.* found similar results and reported heptamers below pH 7 and oligomers consisting out of 13 monomeric units above pH 7. The size of the oligomers was measured using SEC and online DLS and MALS.[91]

IPP5 forms a similar-sized oligomer to Liraglutide with approx. 14 monomer subunits. However, it forms these oligomers at pH 7.5, a pH at which Liraglutide assembles into smaller heptamers. IPP5 oligomers appear to be similar in size to the larger oligomers of Liraglutide detected at high concentrations using static light scattering[227], suggesting that IPP5 and Liraglutide may populate similar species, albeit under different conditions.

Importantly, IPP5, despite having a higher propensity than Liraglutide to form larger oligomers, still assembles further into even bigger aggregates at lower pH values close to and below pH 7. The differences observed between IPP5 and Liraglutide are most likely caused by the difference in net charge or different concentrations at which the measurements were conducted.

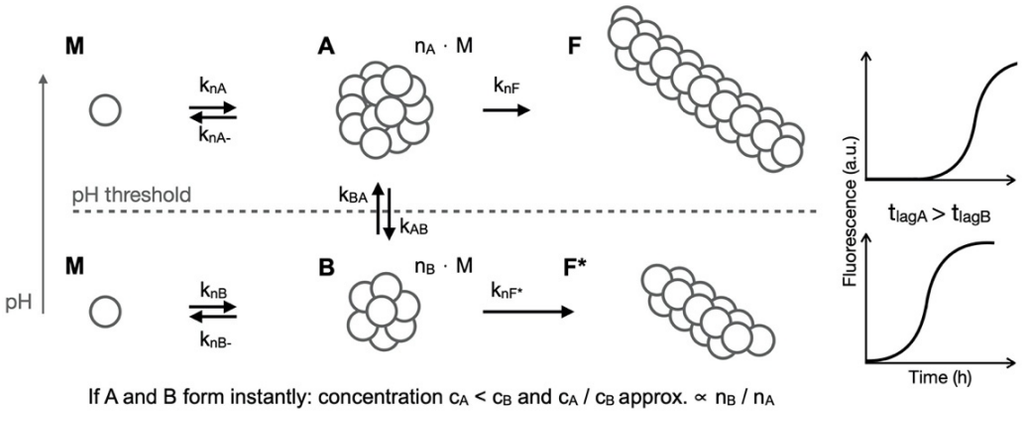


Figure 5.23. Influence of different on-pathway oligomers on the aggregation kinetics.

Different sized on-pathway oligomers at two different pH values associated with different lag times in the aggregation kinetics. F and F* may be fibrils but they may also be some other form of aggregate, for example amorphous aggregate. The process is described by the primary nucleation rate constants k_{nA} and k_{nB}, the dissociation rate constants k_{nA-} and k_{nB-}, the rate constants k_{AB}, k_{BA}, k_{nF} and k_{nF*}, the concentrations of the oligomers c_A and c_B and the oligomer size n_B and n_A.

A comparison of the aggregation propensity of IPP4 and IPP5 using the ThT assay data, Figure 5.7, shows that IPP5, which is initially self-assembled into smaller oligomers, than IPP4, aggregates faster, in agreement with results on

Liraglutide. Bothe *et al*. showed that samples with smaller oligomers (pH 8.1) also aggregate faster.[91] Therefore, it seems that the size and concentration of oligomers may have a direct effect on the aggregation kinetics, independent of lipidation site, C-terminal amidation or pH, Figure 5.23.

5.8.8 Two Different Aggregation Behaviours That Depend on pH

The data for both lipidated peptides, IPP4 and IPP5, show an apparent change IPP4 and IPP5 both show a change in their aggregation kinetics depending on pH, Figure 5.5. A possible explanation for this is that the peptides follow different aggregation pathways to each other, and pathways can vary with pH, Fig. 5.23. For example, IPP4 and IPP5 may populate different-sized oligomers to each other and the size of the oligomers may vary with pH. In support of this hypothesis, a difference in the size of smaller oligomers was measured by DLS for IPP5 between pH 7.0 and 7.5. A similar pH-induced switch between two different oligomeric species was found for Liraglutide by Wang *et al*. and Bothe *et al*. (N.B. Liraglutide is the same as IPP5 but lacks the C-terminal amidation, see Section 5.8.7.[91,227]

The fact that IPP4 and IPP5 both change their aggregation behaviour at a slightly more basic pH value than that observed with Liraglutide can be explained by the lower pI of Liraglutide and subsequent higher negative net charge under these conditions in comparison to IPP4 or IPP5.

5.8.9 Influence of the Amidation and Lipidation on the Secondary Structure

The attachment of a spacer and fatty acid group to Am-GLP-1 affects the secondary structure of the peptide and leads to increase in α-helical structure in freshly dissolved peptide samples. However, amidation at the C-terminus has a similar significant effect on secondary structure than the lipidation itself even though this is a much larger modification, Figure 5.16A. This finding underlines the potent influence of the C-terminus amidation on the secondary structure of GLP-1-like peptides and potentially the role of net and local charge on secondary structure.

5.8.10 Conclusions

One of the most significant influences of lipidation is on peptide solubility. The lipidation restricts the solubility to neutral and basic conditions. This may be due to the change in the net charge of the molecule but is more likely associated with the addition of a large hydrophobic fatty acid moiety. The lipidation also influences the secondary structure. Most notable is the higher fraction of α-helical structure in freshly dissolved samples, however, C-terminal amidation has also been shown to have a similar effect.

The fact that IPP4 shows no formation of fibrils despite having a sigmoidal shaped increase of its ThT fluorescence and an increase of β-structure with time indicates that the ThT signal and change in secondary structure is not always correlated to the formation of amyloid-like fibrils. The ThT signal is, in this case, most likely caused by ThT binding to oligomeric structures or amorphous aggregates or both[362] which still form partial β-sheet structure. Lipidated peptides often form micelles or micelle-like structures due to their amphiphilic nature. For example, for the lipidated analogue of the gut-derived peptide hormone $PYY_{3-36,}$ was shown to forms micelles containing lipopeptides with an α-helical conformation.[34] The amount of β-structure, as measured by far-UV CD and indirectly by ThT binding, leads to the conclusion that if IPP4 forms micelles under these conditions, they must have some β-sheet structure. It can also be concluded that the observed amount of β-structure is not correlated with the intensity of the ThT fluorescence. For example, IPP5 shows a lower ThT fluorescence than IPP4, despite the formation of thin fibrils and the most substantial observed increase of β-structure. This is even more striking due to the fact that IPP5 has a higher negative net charge under basic conditions which should lead to a better binding of positively charged ThT.[291] It would be worth studying the aggregates in more detail by cryo-TEM and SAXS as undertaken for PYY by Hutchinson *et al.* to learn more about these self-assembly structures.

Similar to Liraglutide, a pH-dependent change in aggregation behaviour has been described here for IPP4 and IPP5. In this study, different types of aggregation kinetics were observed dependent upon peptide and pH. A switch between behaviours occurred at pH values approx. 2 units above their pI value.

The different aggregation behaviour is associated with the formation of different sized oligomers or micelle-like structures, these being influenced by the sequence of the peptide as well as the pH conditions.

6 CONCLUSIONS AND FUTURE WORK

6.1 Summary and Conclusions

6.1.1 Background

The work presented in this Thesis is focused on the physical stability and aggregation kinetics of peptides belonging to the GLP-1 family, which is influenced by parameters such as pH, peptide concentration, net charge and chemical modifications (lipidation). Protein/peptide aggregation is a very widely studied topic and draws its importance on one hand from its connection to several human diseases, for example, Alzheimer's, Parkinson's or Huntington Diseases,[37] and on the other hand from issues faced by the pharmaceutical industries who use peptides as biopharmaceutics.[364–366] GLP-1 receptor agonists are under development and used to slow digestion and help lower blood sugar levels.[226,256,262,367] However, the use of therapeutic peptides can be challenging because of problems with their physical stability during long-term storage because of fibrillation and precipitation or their short half-life *in vivo* after injection.[368]

Understanding the formation and proliferation of peptide aggregates under a wide range of conditions is, therefore, of great interest scientifically as well as

commercially.[369] The results of this study have led to a better understanding of the mechanism of aggregation and amyloid formation of GLP-1 and three GLP-1 analogues. GLP-1 is a naturally occurring incretin hormone, and even though it is in extensive use for medicinal purposes, our knowledge about its aggregation kinetics and understanding of this self-assembly process is limited.

A specific focus of these studies was to gain further information on, and understanding of, the unusual behaviour of the aggregation kinetics with respect to its peptide concentration dependence as had been shown previously by the Jackson group.[5] In particular, the aim of the studies presented in this Thesis, was to extend the preliminary studies conducted at pH 7.5-8.2 to acidic pH values and to use time-course experiments to gain further insight into the different species populated during aggregation under different conditions.

In addition, it was crucial to compare the results on the lipidated GLP-1 analogues investigated here with those published for a different lipidated GLP-1 variant, Liraglutide. It had already been shown by Wang and co-workers that changes in pH have significant influence on the aggregation kinetics of Liraglutide.[227] In this case, two different sized oligomers were shown to be populated at different pH values.[227] For the same modified peptide, it was later shown by Bothe *et al*. that an increase in the peptide concentration, lead to longer lag times in aggregation experiments.[91]

6.1.2 Summary of Experimental Chapters

In this Thesis, the first approach was to undertake a more detailed study of GLP-1, Chapter 3, extending the pH range over which the aggregation was monitored. Not only was the aggregation of GLP-1 measured over a wide range of conditions but also at different timepoints to learn more about the mechanism underpinning its aggregation kinetics. In the second study, Chapter 4, the aggregation kinetics of a C-terminal amidated analogue Am-GLP-1 were established and compared to those obtained for GLP-1. The amidation, a rather small perturbation with a predictable influence on the net charge of the peptide, was shown to perturb the aggregation kinetics in a pH-dependent manner. The results enabled the influence of net charge on the aggregation kinetics of GLP-1 and Am-GLP-1 to be quantified. In the final study, the aggregation kinetic of

two lipidated analogues of Am-GLP-1 were investigated and compared with those already obtained for Am-GLP-1 as well as GLP-1, Chapter 5. In all three Chapters, a wide variety of experimental methods were employed. To study the secondary structure of the peptides far-UV circular dichroism (CD), and Fourier-transform infrared spectroscopy (FT-IR) were used. Both datasets were analysed using a singular value decomposition (SVD) method. ANS, tryptophan and *ex-situ* ThT fluorescence were all measured to gain information about the changes in tertiary structure during aggregation. ThT-assays were used to follow the aggregation kinetics and to determine the lag and $t_{1/2}$ of the aggregation process, using a sigmoidal fit of the data. In addition, the ThT-assay data was also fitted with a nucleation-elongation model (NEF) to learn more about the aggregation kinetics and its rate constants. To probe the structure and size of species in solution, dynamic light scattering, atomic force microscopy and scanning electron microscopy were all employed. All experiments were performed over a wide range of conditions.

6.1.3 Symmetrical Aggregation Propensity Around the pI

Results of measurements of the aggregation of GLP-1 under acidic conditions, Chapter 3, showed that a relatively small change in pH altered the aggregation behaviour, illustrating that the self-assembly of GLP-1 is very sensitive to pH not only in basic conditions but also in the acidic pH regime. Results showed that at pH 3.5 but not pH 3.0, off-pathway oligomers become populated leading to a change in the dependence of the aggregation kinetics on peptide concentration. The results demonstrate that the aggregation propensity and the population of off-pathway species is strongly related to the pH and the pI of the peptides. This was also shown to be true for the C-terminally amidated variant of GLP-1, Am-GLP-1. The population of off-pathway oligomers for GLP-1 and AM-GLP-1 surprisingly results in relatively long lag times close to their pIs and therefore also under conditions of relatively low net charge on the peptides. This is stark contrast to previous studies on other peptides where they were shown to aggregate faster if they carried a lower net charge.[131–136] However, the findings for GLP-1 and its analogues can be explained by the fact that they have a significant propensity to form off-pathway species near their

pI values, whilst other peptides, such as Aβ, do not.[85,370]

6.1.4 The Net Charge Influences the Aggregation Propensity more than the Secondary Structure of the Monomer

The motivation to study the aggregation of Am-GLP-1 and to compare it to that of GLP-1 came from two ideas: First, it enables the influence of the net charge of the peptide on its aggregation propensity to be determined without a significant change in the sequence and, secondly, it provides a starting point for further studies to assess the impact of chemical modification (lipidation) on physical stability. This is because the two lipidated forms of GLP-1 available are both amidated at the C-terminus. The results clearly show, for all four peptides, several properties of the peptide are correlated with net charge.

Starting with the protonation sites (and state) of the peptides, these directly affect the net charge but were also shown to influence the secondary structure of the monomeric peptide as well as the pI and similarly the solubility of the peptide at different pH values. Net charge and/or secondary structure both affect the aggregation properties. Results described in Chapters 3 and 4, strongly suggest that the net charge has a more significant influence on the aggregation kinetics compared with secondary structure, Section 4.13.2.

6.1.5 Changes of Secondary Structure at pH Values near the pK_a Values of Specific Side Chains Influence Aggregation Propensity

The results of the studies described in Chapters 3 and 4 also suggest that the secondary structure also influences the aggregation propensity. However, the relationship between the secondary structure and the aggregation propensity is more specific to certain conditions. The far-UV CD data suggests that, especially in the acidic regime between pH 3.0 and 4.5, the de-/protonation of the side chains of aspartic acid and glutamic acid cause a significant increase of the α-helical structure of both non-lipidated peptides, Figure 4.32A. The observed change in secondary structure most likely affects the aggregation propensity of both peptides similarly. While at pH 3.0, both peptides follow

nucleation-elongation kinetics with decreasing lag times with increasing peptide concentration, leads the increase in alpha-helical structure to a change of the aggregation kinetics at less acidic pH values. For example, Am-GLP-1 shows constant lag time with increasing peptide concentration at pH 4.0 and GLP-1 shows even increasing lag time with increasing peptide concentration already at pH 3.5. Both behaviours are typical for the formation of off-pathway species. The influence of the secondary structure on the aggregation propensity under these conditions may also explain why, for example, GLP-1 has similar aggregation propensities in the acidic and basic pH regimes despite having very different net charges.

Notable, and against our expectations, a substantial effect of the amidation on the secondary structure of the peptide was observed under neutral and basic condition: Am-GLP-1 having more α-helical structure than GLP-1, over wide range of conditions. It is also of note that, under all the conditions at which the aggregation kinetics follow simple nucleation-elongation kinetics where only on-pathway oligomers are significantly populated, the peptides show a higher fraction of β-sheet in comparison to α-helical structure in the monomeric state.

In summary, the data suggests the net charge is the main factor influencing the aggregation kinetics, however, under specific conditions, there is an additional effect and the secondary structure in the monomer also plays a role.

6.1.6 Lipidation and Amidation Increase α-Helicity of Freshly Dissolved Peptides

The lipidation of peptides is a widely used modification in peptide therapeutics which improve half-lives *in vivo* as well as increasing the physical stability of a peptide *in vitro*. The latter is significant in the improvement of drug formulations.[3,371]

The study of IPP4 and IPP5 described in Chapter 5 shows that the solubility of GLP-1 is greatly affected by lipidation resulting in them only being soluble at neutral/basic pH values above their pIs. It was previously reported that lipidation increases the α-helical content of freshly dissolved samples in a study of Liraglutide, a lipidated GLP-1 analogue.[224] The data presented in Chapter 5

confirms this finding and extends the conclusion to other lipidated GLP-1 analogues. However, by comparison of the data on IPP4 and IPP5 with GLP-1 and Am-GLP-1, it becomes clear that C-terminal amidation itself is responsible for the increase in α-helicity observed.

6.1.7 Lipidated Peptides Self-Assemble to form β-Sheet Structures Independent of the Formation of Rigid Fibrils

SEM imaging of the lipidated peptides which had been incubated under the same conditions as those that induced amyloid fibrils in GLP-1 and Am-GLP-1, showed that they have a much lower tendency to form rigid fibrils. At the same time, the ThT fluorescence and the far-UV CD data showed the formation of β-sheet during their aggregation. It can therefore be concluded that these two lipidated variants of GLP-1 form a β-sheet rich oligomeric structure, which does not appear to be able to convert into amyloid-like fibrils over the timescales used in this study.

6.1.8 GLP-1 Forms β-Sheet Rich Intermediate Species

The time-course measurements of the secondary structure of GLP-1 under basic conditions showed that GLP-1 forms intermediate oligomers/aggregates that contain considerable β-sheet content during the lag phase. Surprisingly, these aggregates cannot bind to ThT. There is some evidence of short fibril-like aggregates already present during the lag phase, as shown by AFM in Figures 3.17D and 3.18D, however, it is likely that these species, which are likely to bind to ThT, are not present in significant amounts. This suggests that fibrillar the oligomers/aggregates formed which contribute to the increase in β-sheet signal are globular in nature.

6.1.9 pH Switch Observed for Four Different GLP-1 Analogues

The lipidated variants IPP4 and IPP5 also show pH-dependent self-assembly behaviour similar to GLP-1 and Am-GLP-1. GLP-1 changes its aggregation behaviour twice, between pH 3.0 and pH 3.5 as well as between pH 7.5 and pH 8.2. Similarly, Am-GLP-1 changes its aggregation kinetics between pH 3.0 and

4.0. Also, in the basic regime, a pH switch can be observed for Am-GLP-1, however, the pH value at which this occurs is not as well defined as in the case of GLP-1. Measurements on the two lipidated peptides were restricted to the basic regime where they are soluble, but they also showed a pH-dependent aggregation behaviour. IPP4 changes its aggregation behaviour between pH 7.7 and 8.2 and IPP5 between pH 7.5 and pH 7.7. These results indicate that the pH dependence is a common feature of the aggregation of all GLP-1 analogues. The data strongly suggests that the pH value at which the self-assembly behaviour changes is related to the pI of the peptides.

6.1.10 Physical Evidence for Different Aggregates Formed Under Different pH Conditions for Individual Peptides

Even though GLP-1 and Am-GLP-1 form similar fibrillar species over a wide range of pH values, there is evidence that the structure and aggregation pathways can vary. For example, GLP-1 showed different sized curly fibrils during the lag phase under acidic and basic conditions. For Am-GLP-1, variations in high-order structures were observed under neutral and basic conditions, indicating indirectly that the fibrillar structures formed might be different. Finally, both peptides also show changes in their ANS binding affinity between acidic and basic conditions. Similarly, the Trp side chain in the fibrils are more buried and has less contact with the aqueous solvent under acidic conditions.

6.1.11 Rate Constants are not Globally Valid Under Different Sample Concentrations

In the last decade, much work has gone into developing methods for the fitting of kinetic data from aggregation reactions, simulation of aggregation and prediction of peptide aggregation propensity. From the fitting of experimental aggregation data for some peptides to different mathematical models, it has been possible, in some cases, to identify some of the critical processes underpinning the self-assembly reactions on the molecular level. These mathematical models have helped to generate information on the relative importance of primary nucleation, elongation, fragmentation and surface-

catalysed secondary nucleation.[83,353,354] Recently, the active suppression of fibril elongation by off-pathway oligomers was shown for human and salmon calcitonin (hCT/sCT).[90]

In this study, attempts to fit the experimental data for the aggregation of GLP-1 and Am-GLP-1 at different peptide concentrations to a NEF model (nucleation, elongation and fragmentation) were undertaken. However, it was not possible to fit the data for different peptide concentrations globally using this model and these methods, which shows that, in this case, the rate constants of the NEF model are not independent of the peptide concentration. This is almost certainly because there are off-pathways species populated during the aggregation of GLP-1 and Am-GLP-1 requiring a more complex model and fitting algorithm to be developed.

6.2 Further Work

6.2.1 Further Characterisation of the Structure of Oligomeric Species

To increase our understanding of the pH dependence of GLP-1/Am-GLP-1 aggregation, more detailed measurements on the size and structure of the oligomeric species populated at different pH values is essential. For some of the peptides/conditions studied, DLS and SEC were used to measure the size of the oligomers. Additional measurements using both techniques, for all four peptides over a wider range of conditions, especially close to and around the pH conditions at which all four peptides change their aggregation behaviour, would be beneficial. In addition, DLS experiments which were undertaken at different time points for GLP-1, showed a high variability, Figure 3.25A-D, likely because the concentration of peptide was insufficient. The DLS experiments should be repeated at higher peptide concentrations, and ThT aggregation assays should be performed under the same conditions to allow a true comparison of the different datasets. For a more detailed analysis of the molecular weight of the oligomers formed during aggregation, alternative techniques such as SAXS[363] and SEC-MALS[91,372] could be employed.

Another promising approach to investigate the aggregation kinetics of peptides and small proteins is the combination of mass spectrometry (MS) and its allied technique of ion mobility mass spectrometry (IM-MS). IMS provides information on the cross-sectional area of ions in the gas phase as measured by their mobility through a buffer gas, as well as providing information on the mass of the corresponding species.[373] IM-MS can therefore separate complex mixtures and ions which are indistinguishable by mass spectrometry alone. For excellent reviews on how IM-MS is used to study aggregation and amyloid formation, reporting recommendations for IM-MS data and the different IM-MS techniques used in drift-time ion mobility spectrometry (DTIMS), traveling-wave ion mobility spectrometry (TWIMS) and field-asymmetry ion mobility spectrometry (FAIMS) see Ashcroft (2010)[374], Gabelica *et al.* (2019)[375] and Lanucara *et al*. (2014)[376]

To analyse the aggregation pathway of peptides and proteins using MS, and specifically ESI-MS (electrospray ionisation-MS) or nano-ES-MS (nanoflow electrospray-MS), it is crucial that non-covalently bound peptide complexes retain, at least to a certain extent, their tertiary and quaternary structure during ionisation and in the gas-phase. This has been shown to be possible in a number of cases, for example, in measurements of insulin aggregation conducted by Nettleton *et al*.[377] Their nano-ES-MS data showed that insulin in the presence of Zn^{2+} is present in a mix of dimers, tetramers and hexamers.[377] The latter with two to four Zn^{2+} ions bound. The absence of small odd-numbered oligomers in the gas phase was in good agreement with previous measurements in solution by Lord *et al.*[378] which together with the findings of Nettleton *et al*. provided evidence that no additional oligomers form as an artefact of the electrospray ionisation or the fact that the experiment is conducted in the gas phase.[377]

Bernstein and co-workers used MS techniques to compare the oligomers formed by $A\beta_{1-42}$ with oligomers formed by $A\beta_{1-40}$.[379,380] Such a comparison is of interest because their propensity to aggregate into amyloid fibrils is very different. $A\beta_{1-42}$ showed evidence for the formation of dimeric, tetrameric, hexameric and dodecameric oligomers, whereas $A\beta_{1-40}$ showed only evidence of dimer and tetramer.[379,380]

IM-MS was also successfully used in a study of $A\beta_{10-30}$. This fragment, containing the Aβ amyloid core region, showed the presence of monomers, dimers, trimers, tetramers, pentamers and hexamers.[381] To gain more information about the structure, Jablonowska *et al*. also combined their IM-MS experiments with HDX.[381] The results indicated no difference in the rate of exchange for different oligomers suggesting that the oligomers lack a defined or stable structure or that the exchange of monomeric subunits is fast. A comparison with ^{15}N-labeled peptides confirmed the latter and showed evidence of an exchange of monomers with subunits in the oligomers on a sub-second time scale.[381]

Together, these results show that IM-MS can be a powerful method to study the aggregation kinetics and the formation of oligomers arising from monomers. However, IM-MS results have to be considered carefully. It is, for example, not immediately given that every detected species in the gas phase is also naturally occurring in solution and not merely an artefact of the ionisation technique. It is also possible that different species are suppressed. This question is closely related to the general challenge of proving that fibril formation *in vivo* can be reproduced *in vitro*. Both can only be solved by carefully comparing the results of many different biophysical techniques.

6.2.2 Further Characterisation of the Atomic Structure of Fibrils under Different pH Conditions

The studies presented in this Thesis showed some evidence that GLP-1 and Am-GLP-1 form different aggregates under different pH conditions, Section 6.1.10. Additional, solid-state NMR measurements could resolve if the differences observed in the aggregates are caused by altered atomic structures and conformations of the β-strands forming the β-sheets. This would also enable a comparison to published results on the fibrils of glucagon formed under acidic conditions, in which the entire peptide length is involved in cross-β bonding.[65]

6.2.3 Is the Aggregation of GLP-1 and GLP-1 Analogues Influenced by Liquid-Liquid Phase Transition or Gelation?

IM-MS methods discussed in Section 6.2.1 can be described as bottom-up approaches measuring the initial steps in aggregation, which consequently influence the aggregation behaviour. Alternatively, a detailed study of the bulk behaviour of dissolved peptides and therefore, a top-down approach could also lead to valuable insights. A broader comparison between the ThT fluorescence and the change of the turbidity or viscosity of the samples could provide information on whether the system undergoes a liquid-liquid phase transition or gelation is happening. A change of these parameters would raise the question of whether they influence the aggregation kinetics or not.

Wang *et al.* described a liquid-liquid phase transition (LLPS) for six palmitoylated as well as one unconjugated incretin peptides.[382] For the lipopeptide, IP5 which has a pI of 4.0 and molecular weight of 4500 Da, quasi-elastic light-scattering was measured at several pH values between pH 5.5 and 7.42. The peptide solution appeared either in a clear or cloudy phase depending on the temperature, which could be changed reversibly by lowering or increasing the temperature. The data showed that the temperature at which phase separation occurred decreased with increasing pH and therefore increasing net charge on the peptide. The peptide concentrations used were very high, starting at 2250 μM, and were at least a magnitude higher in comparison to the concentrations used in the studies undertaken and reported in this Thesis.

Gelation caused by peptides and especially peptide amphiphiles has been previously reported.[34,72,383–386] The preferred secondary structure of peptides forming hydrogels is the β-sheet structure,[386] however, hydrogels can also form from peptides rich in α-helical structure.[387] A detailed study of the potential gelation of GLP-1 and GLP-1 analogues samples could clarify if gelation occurs under the conditions used in these studies.

6.2.4 Complementary Experiments for a Deeper Understanding of the Effect of Lipidation

One primary aim of this study was to learn more about the influence of lipidation on the physical stability of GLP-1. Am-GLP-1 was studied in detail to be able to compare its aggregation kinetics with the kinetics of the lipidated peptides which are both analogues of Am-GLP-1. However, both lipidated peptides had additional modifications of their sequence to remove lysine side chains to avoid additional covalent attachment of the spacer and fatty acid. In these cases, a lysine was mutated to arginine which retains the positive charge. It is assumed that these mutations have a rather small effect on the aggregation behaviour. However, it was shown in a separate study that substitution of lysine with arginine can influence a peptide's aggregation kinetics.[388] Therefore, it would be valuable to also measure the sequences of Am-GLP-1 which correspond directly to the non-lipidated form of both IPP4 and IPP5. This would exclude any misinterpretation of characteristics observed for the lipidated peptides that were due to the lysing-arginine substitution.

In addition, a detailed study of the effect of different linker types as well as different fatty acids, could lead to valuable information regarding how the exact nature of the lipidation modifications affect physical stability. This would enable an assessment of the contribution of different parts of the lipidated GLP-1 variants to physical stability and whether the properties of the different regions of the lipidated peptides acted in an additive fashion or if the product of peptide, linker and fatty acid show independent characteristics.

6.2.5 Role of Zeta Potential in the Aggregation of GLP-1 and GLP-1 Analogues

The results in Chapter 4 show the influence of the net charge on the aggregate kinetics. A detailed study of the zeta potential, especially at the pH values summarised in Section 6.1.9 could help to understand the influence of the net charge on off-pathway oligomers or aggregates. It might also help to understand why under certain conditions samples with lower net charge are more likely to form off-pathway species and show longer aggregation lag times.

Learning more about the zeta potential is also valuable for the understanding of the binding affinity of fluorescent molecules such as ANS or ThT. It was recently shown that the zeta potential is related to the binding and fluorescence of thioflavin T.[291] In this study, it was shown that the zeta potential rather than pH-dependent assembly of the fibrils is the primary factor affecting ThT binding and fluorescence. Interestingly, the ThT fluorescence measurements of GLP-1 and GLP-1 analogues under acidic and basic condition show similar maximum fluorescence intensities. This means that either the assumption of a significant influence of the zeta potential is not correct for GLP-1 or the fibrils formed under acidic conditions are structurally different in comparison to fibrils formed under basic conditions and the influence of the different structures on the ThT fluorescence is more dominant than the influence of the zeta potential.

7 REFERENCES

1. Fosgerau, K. and Hoffmann, T. (2015). Peptide therapeutics: current status and future directions. Drug Discovery Today *20*, 122–128.

2. Lau, J.L. and Dunn, M.K. (2018). Therapeutic peptides: Historical perspectives, current development trends, and future directions. Bioorganic & Medicinal Chemistry *26*, 2700–2707.

3. Frokjaer, S. and Otzen, D.E. (2005). Protein drug stability: a formulation challenge. Nature Reviews Drug Discovery *4*, 298–306.

4. Moussa, E.M., Panchal, J.P., Moorthy, B.S., Blum, J.S., Joubert, M.K., Narhi, L.O. and Topp, E.M. (2016). Immunogenicity of therapeutic protein Aggregates. Journal of Pharmaceutical Sciences *105*, 417–430.

5. Zapadka, K.L., Becher, F.J., Uddin, S., Varley, P.G., Bishop, S., Santos, A.L.G.D. and Jackson, S.E. (2016). A pH-Induced Switch in Human Glucagon-like Peptide-1 Aggregation Kinetics. Journal of the American Chemical Society *138*, 16259–16265.

6. Poon, S., Birkett, N., Fowler, S., Luisi, B., Dobson, C. and Zurdo, J. (2009). Amyloidogenicity and Aggregate Cytotoxicity of Human Glucagon-Like Peptide-1 (hGLP-1). Protein & Peptide Letters *16*, 1548–1556.

7. Pedersen, J.S. (2010). The Nature of Amyloid-like Glucagon Fibrils. Journal of Diabetes Science and Technology *4*, 1357–1367.

8. Pedersen, J.S., Andersen, C.B. and Otzen, D.E. (2010). Amyloid structure - one but not the same: the many levels of fibrillar polymorphism. FEBS Journal *277*, 4591–4601.

9. Jong, K.L.D., Incledon, B., Yip, C.M. and DeFelippis, M.R. (2006). Amyloid Fibrils of Glucagon Characterized by High-Resolution Atomic Force Microscopy. Biophysical Journal *91*, 1905–1914.

10. Caputo, N., Jackson, M.A., Castle, J.R., Youssef, J.E., Bakhtiani, P.A., Bergstrom, C.P., Carroll, J.M., Breen, M.E., Leonard, G.L., David, L.L., Roberts, C.T. and Ward, W.K. (2014). Biochemical Stabilization of Glucagon at Alkaline pH. Diabetes Technology & Therapeutics *16*, 747–758.

11. Caputo, N., Castle, J.R., Bergstrom, C.P., Carroll, J.M., Bakhtiani, P.A., Jackson, M.A., Roberts, C.T., David, L.L. and Ward, W.K. (2013). Mechanisms of glucagon degradation at alkaline pH. Peptides *45*, 40–47.

12. Gilchrist, P.J. and Bradshaw, J.P. (1993). Amyloid formation by salmon calcitonin. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease *1182*, 111–114.

13. Gaudiano, M.C., Colone, M., Bombelli, C., Chistolini, P., Valvo, L. and Diociaiuti, M. (2005). Early stages of salmon calcitonin aggregation: Effect induced by ageing and oxidation processes in water and in the presence of model membranes. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1750*, 134–145.

14. Diociaiuti, M., Macchia, G., Paradisi, S., Frank, C., Camerini, S., Chistolini, P., Gaudiano, M.C., Petrucci, T.C. and Malchiodi-Albedi, F. (2014). Native metastable prefibrillar oligomers are the most neurotoxic species among amyloid aggregates. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease *1842*, 1622–1629.

15. Diociaiuti, M., Gaudiano, M.C. and Malchiodi-Albedi, F. (2011). The Slowly Aggregating Salmon Calcitonin: A Useful Tool for the Study of the Amyloid Oligomers Structure and Activity. International Journal of Molecular Sciences *12*, 9277–9295.

16. Makin, O.S. and Serpell, L.C. (2004). Structural Characterisation of Islet Amyloid Polypeptide Fibrils. Journal of Molecular Biology *335*, 1279–1288.

17. Luca, S., Yau, W.-M., Leapman, R. and Tycko, R. (2007). Peptide Conformation and Supramolecular Organization in Amylin Fibrils: Constraints from Solid-State NMR. Biochemistry *46*, 13505–13522.

18. Bedrood, S., Li, Y., Isas, J.M., Hegde, B.G., Baxa, U., Haworth, I.S. and Langen, R. (2011). Fibril Structure of Human Islet Amyloid Polypeptide. Journal of Biological Chemistry *287*, 5235–5241.

19. Jha, S., Snell, J.M., Sheftic, S.R., Patil, S.M., Daniels, S.B., Kolling, F.W. and Alexandrescu, A.T. (2014). pH Dependence of Amylin Fibrillization. Biochemistry *53*, 300–310.

20. Weirich, F., Gremer, L., Mirecka, E.A., Schiefer, S., Hoyer, W. and Heise, H. (2016). Structural Characterization of Fibrils from Recombinant Human

Islet Amyloid Polypeptide by Solid-State NMR: The Central FGAILS Segment Is Part of the β-Sheet Core. PLoS ONE *11*, e0161243.

21. Jimenez, J.L., Nettleton, E.J., Bouchard, M., Robinson, C.V., Dobson, C.M. and Saibil, H.R. (2002). The protofilament structure of insulin amyloid fibrils. Proceedings of the National Academy of Sciences *99*, 9196–9201.

22. Hua, Q. and Weiss, M.A. (2004). Mechanism of Insulin Fibrillation. Journal of Biological Chemistry *279*, 21449–21460.

23. Smith, J.F., Knowles, T.P.J., Dobson, C.M., MacPhee, C.E. and Welland, M.E. (2006). Characterization of the nanoscale properties of individual amyloid fibrils. Proceedings of the National Academy of Sciences *103*, 15806–15811.

24. Kurouski, D., Deckert-Gaudig, T., Deckert, V. and Lednev, I.K. (2012). Structure and Composition of Insulin Fibril Surfaces Probed by TERS. Journal of the American Chemical Society *134*, 13323–13329.

25. Kurouski, D., Dukor, R.K., Lu, X., Nafie, L.A. and Lednev, I.K. (2012). Normal and Reversed Supramolecular Chirality of Insulin Fibrils Probed by Vibrational Circular Dichroism at the Protofilament Level of Fibril Structure. Biophysical Journal *103*, 522–531.

26. Kurouski, D., Dukor, R.K., Lu, X., Nafie, L.A. and Lednev, I.K. (2012). Spontaneous inter-conversion of insulin fibril chirality. Chemical Communications *48*, 2837.

27. Kurouski, D., Washington, J., Ozbil, M., Prabhakar, R., Shekhtman, A. and Lednev, I.K. (2012). Disulfide Bridges Remain Intact while Native Insulin Converts into Amyloid Fibrils. PLoS ONE *7*, e36989.

28. Yoshihara, H., Saito, J., Tanabe, A., Amada, T., Asakura, T., Kitagawa, K. and Asada, S. (2016). Characterization of Novel Insulin Fibrils That Show Strong Cytotoxicity Under Physiological pH. Journal of Pharmaceutical Sciences *105*, 1419–1426.

29. Smith, M.I., Foderà, V., Sharp, J.S., Roberts, C.J. and Donald, A.M. (2012). Factors affecting the formation of insulin amyloid spherulites. Colloids and Surfaces B: Biointerfaces *89*, 216–222.

30. Kerr, D., Wizemann, E., Senstius, J., Zacho, M. and Ampudia-Blasco, F.J. (2013). Stability and Performance of Rapid-Acting Insulin Analogs Used for Continuous Subcutaneous Insulin Infusion: A Systematic Review. Journal of Diabetes Science and Technology *7*, 1595–1606.

31. Gopalswamy, M., Kumar, A., Adler, J., Baumann, M., Henze, M., Kumar, S.T., Fändrich, M., Scheidt, H.A., Huster, D. and Balbach, J. (2015). Structural characterization of amyloid fibrils from the human parathyroid hormone. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1854*, 249–257.

32. Hermeling, S., Aranha, L., Damen, J.M.A., Slijper, M., Schellekens, H., Crommelin, D.J.A. and Jiskoot, W. (2005). Structural Characterization and Immunogenicity in Wild-Type and Immune Tolerant Mice of Degraded Recombinant Human Interferon Alpha2b. Pharmaceutical Research *22*, 1997–2006.

33. Torosantucci, R., Sharov, V.S., Beers, M. van, Brinks, V., Schöneich, C. and Jiskoot, W. (2013). Identification of Oxidation Sites and Covalent Cross-Links in Metal Catalyzed Oxidized Interferon Beta-1a: Potential Implications for Protein Aggregation and Immunogenicity. Molecular Pharmaceutics *10*, 2311–2322.

34. Hutchinson, J.A., Burholt, S., Hamley, I.W., Lundback, A.-K., Uddin, S., Santos, A.G. dos, Reza, M., Seitsonen, J. and Ruokolainen, J. (2018). The Effect of Lipidation on the Self-Assembly of the Gut-Derived Peptide Hormone PYY3–36. Bioconjugate Chemistry *29*, 2296–2308.

35. Lipiäinen, T., Peltoniemi, M., Sarkhel, S., Yrjönen, T., Vuorela, H., Urtti, A. and Juppo, A. (2015). Formulation and Stability of Cytokine Therapeutics. Journal of Pharmaceutical Sciences *104*, 307–326.

36. Krishnan, S., Chi, E.Y., Webb, J.N., Chang, B.S., Shan, D., Goldenberg, M., Manning, M.C., Randolph, T.W. and Carpenter, J.F. (2002). Aggregation of Granulocyte Colony Stimulating Factor under Physiological Conditions: Characterization and Thermodynamic Inhibition. Biochemistry *41*, 6422–6431.

37. Knowles, T.P., Vendruscolo, M. and Dobson, C.M. (2014). The amyloid state and its association with protein misfolding diseases. Nature Reviews Molecular Cell Biology *15*, 384–396.

38. Hamley, I.W. (2012). The Amyloid Beta Peptide: A Chemist's Perspective. Role in Alzheimer's and Fibrillization. Chemical Reviews *112*, 5147–5192.

39. Bennett, M.C. (2005). The role of α-synuclein in neurodegenerative diseases. Pharmacology & Therapeutics *105*, 311–331.

40. Delenclos, M., Burgess, J.D., Lamprokostopoulou, A., Outeiro, T.F., Vekrellis, K. and McLean, P.J. (2019). Cellular models of alpha-synuclein toxicity and aggregation. Journal of Neurochemistry *150*, 566–576.

41. Jansen, A.H.P., Batenburg, K.L., Pecho-Vrieseling, E. and Reits, E.A. (2017). Visualization of prion-like transfer in Huntington's disease models. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease *1863*, 793–800.

42. Kyle, R.A. (2001). Amyloidosis: A Convoluted Story. British Journal of Haematology, 529–538.

43. Virchow, R. (1989). As Based upon Physiological and Pathological Histology. Nutrition Reviews *47*, 23–25.

44. Aterman, K. (1976). A historical note on the iodine-sulphuric acid reaction of amyloid. Histochemistry *49*, 131–143.

45. Yakupova, E.I., Bobyleva, L.G., Vikhlyantsev, I.M. and Bobylev, A.G. (2019). Congo Red and amyloids: history and relationship. Bioscience Reports *39*.

46. Shirahama, T. and Cohen, A.S. (1965). Structure of Amyloid Fibrils after Negative Staining and High-resolution Electron Microscopy. Nature *206*, 737–738.

47. Astbury, W.T. and Street, A. (1931). X-ray studies of the structure of hair, wool, and related fibres.- I. General. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character *230*, 75–101.

48. Astbury, W.T. and Woods, H.J. (1933). X-rays studies of the structure of hair, wool, and related fibres. II. - The molecular structure and elastic properties of hair keratin. Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character *114*, 314–316.

49. Astbury, W.T. (1933). Some problems in the X-ray analysis of the structure of animal hairs and other protein fibres. Transactions of the Faraday Society *29*, 193–205.

50. Astbury, W.T. and Lomax, R. (1934). X-Ray Photographs of Crystalline Pepsin. Nature *133*, 795.

51. Cohen, A.S. and Calkins, E. (1959). Electron Microscopic Observations on a Fibrous Component in Amyloid of Diverse Origins. Nature *183*, 1202–1203.

52. Shirahama, T. and Cohen, A.S. (1967). High-Resolution Electron Microscopic Analysis of the Amyloid Fibril. The Journal of Cell Biology *33*, 679–708.

53. Geddes, A.J., Parker, K.D., Atkins, E.D.T. and Beighton, E. (1968). "Cross-β" conformation in proteins. Journal of Molecular Biology *32*, 343–358.

54. Eanes, E.D. and Glenner, G.G. (1968). X-ray diffraction studies on amyloid filaments. Journal of Histochemistry and Cytochemistry *16*, 673–677.

55. Ke, P.C., Zhou, R., Serpell, L.C., Riek, R., Knowles, T.P.J., Lashuel, H.A., Gazit, E., Hamley, I.W., Davis, T.P., Fändrich, M., Otzen, D.E., Chapman, M.R., Dobson, C.M., Eisenberg, D.S. and Mezzenga, R. (2020). Half a century of amyloids: past, present and future. Chemical Society Reviews *49*, 5473–5509.

56. Eisenberg, D. and Jucker, M. (2012). The Amyloid State of Proteins in Human Diseases. Cell *148*, 1188–1203.

57. Fitzpatrick, A.W.P., Debelouchina, G.T., Bayro, M.J., Clare, D.K., Caporini, M.A., Bajaj, V.S., Jaroniec, C.P., Wang, L., Ladizhansky, V., Müller, S.A., … Dobson, C.M. (2013). Atomic structure and hierarchical assembly of a cross-β amyloid fibril. Proceedings of the National Academy of Sciences *110*, 5468–5473.

58. Riek, R. and Eisenberg, D.S. (2016). The activities of amyloids from a structural perspective. Nature *539*, 227–235.

59. Eisenberg, D.S. and Sawaya, M.R. (2016). Structural Studies of Amyloid Proteins at the Molecular Level. Annual Review of Biochemistry *86*, 1–27.

60. Tabor, R.F., Grieser, F., Dagastine, R.R. and Chan, D.Y.C. (2014). The hydrophobic force: measurements and methods. Physical Chemistry Chemical Physics *16*, 18065–18075.

61. Collinge, J. and Clarke, A.R. (2007). A general model of prion strains and their pathogenicity. Science *318*, 930–936.

62. Morales, R., Abid, K. and Soto, C. (2007). The prion strain phenomenon: molecular basis and unprecedented features. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease *1772*, 681–691.

63. Hamley, I.W. (2007). Peptide Fibrillization. Angewandte Chemie International Edition *46*, 8128–8147.

64. Wei, G., Su, Z., Reynolds, N.P., Arosio, P., Hamley, I.W., Gazit, E. and Mezzenga, R. (2017). Self-assembling peptide and protein amyloids: from structure to tailored function in nanotechnology. Chemical Society Reviews *46*, 4661–4708.

65. Gelenter, M.D., Smith, K.J., Liao, S.-Y., Mandala, V.S., Dregni, A.J., Lamm, M.S., Tian, Y., Xu, W., Pochan, D.J., Tucker, T.J., Su, Y. and Hong, M. (2019). The peptide hormone glucagon forms amyloid fibrils with two coexisting β-strand conformations. Nature Structural & Molecular Biology *26*, 592–598.

66. Hiller, S., Garces, R.G., Malia, T.J., Orekhov, V.Y., Colombini, M. and Wagner, G. (2008). Solution Structure of the Integral Human Membrane Protein VDAC-1 in Detergent Micelles. Science *321*, 1206–1210.

67. Petkova, A.T., Yau, W.-M. and Tycko, R. (2006). Experimental Constraints on Quaternary Structure in Alzheimer's β-Amyloid Fibrils. Biochemistry *45*, 498–512.

68. Pedersen, J.S., Dikov, D., Flink, J.L., Hjuler, H.A., Christiansen, G. and Otzen, D.E. (2006). The Changing Face of Glucagon Fibrillation: Structural Polymorphism and Conformational Imprinting. Journal of Molecular Biology *355*, 501–523.

69. Krebs, M.R.H., Domike, K.R. and Donald, A.M. (2009). Protein aggregation: more than just fibrils. Biochemical Society Transactions *37*, 682–686.

70. Bowerman, C.J. and Nilsson, B.L. (2012). Review self-assembly of amphipathic β-sheet peptides: Insights and applications. Biopolymers *98*, 169–184.

71. Pelin, J.N.B.D., Gerbelli, B.B., Edwards-Gayle, C.J.C., Aguilar, A.M., Castelletto, V., Hamley, I.W. and Alves, W.A. (2020). Amyloid Peptide Mixtures: Self-Assembly, Hydrogelation, Nematic Ordering, and Catalysts in Aldol Reactions. Langmuir *36*, 2767–2774.

72. Djabourov, M., Nishinari, K. and Ross-Murphy, S.B. (2013). 9 - Colloidal gels from proteins and peptides. In Physical Gels from Biological and Synthetic Polymers Physical Gels from Biological and Synthetic Polymers., 256–286.

73. Trainor, K., Broom, A. and Meiering, E.M. (2017). Exploring the relationships between protein sequence, structure and solubility. Current Opinion in Structural Biology *42*, 136–146.

74. Argudo, P.G. and Giner-Casares, J.J. (2021). Folding and self-assembly of short intrinsically disordered peptides and protein regions. Nanoscale Advances *3*, 1789–1812.

75. Jahn, T.R. and Radford, S.E. (2008). Folding versus aggregation: Polypeptide conformations on competing pathways. Archives of Biochemistry and Biophysics *469*, 100–117.

76. Uemura, E., Niwa, T., Minami, S., Takemoto, K., Fukuchi, S., Machida, K., Imataka, H., Ueda, T., Ota, M. and Taguchi, H. (2018). Large-scale aggregation analysis of eukaryotic proteins reveals an involvement of intrinsically disordered regions in protein folding. Scientific Reports *8*, 678.

77. Levine, Z.A., Larini, L., LaPointe, N.E., Feinstein, S.C. and Shea, J.-E. (2015). Regulation and aggregation of intrinsically disordered peptides. Proceedings of the National Academy of Sciences *112*, 2758–2763.

78. Come, J.H., Fraser, P.E. and Landsbury, P.T. (1993). A kinetic model for amyloid formation in the prion diseases: importance of seeding. Proceedings of the National Academy of Sciences *90*, 5959–5963.

79. Jarrett, J.T. and Lansbury, P.T. (1993). Seeding "One-Dimensional Crystallization" of Amyloid: A Pathogenic Mechanism in Alzheimer's Disease and Scrapie? Cell *73*, 1055–1058.

80. Zimm, B.H. and Bragg, J.K. (1959). Theory of the phase transition between helix and random coil in polypeptide chains. Journal of Chemical Physics *31*, 526–535.

81. Oosawa, F. and Kasai, M. (1962). A theory of linear and helical aggregations of macromolecules. Journal of Molecular Biology *4*, 10–21.

82. Eaton, W.A., Ferrone, F.A. and Hofrichter, J. (1985). Kinetics of sickle hemoglobin polymerization: II. A double nucleation mechanism. Journal of molecular biology *183*, 611–631.

83. Cohen, S.I.A., Vendruscolo, M., Welland, M.E., Dobson, C.M., Terentjev, E.M. and Knowles, T.P.J. (2011). Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments. The Journal of Chemical Physics *135*.

84. Meisl, G., Yang, X., Dobson, C.M., Linse, S. and Knowles, T.P. (2017). Modulation of electrostatic interactions to reveal a reaction network unifying the aggregation behaviour of the Aβ42 peptide and its variants. Chemical science *8*, 4352–4362.

85. Cohen, S.I., Linse, S., Luheshi, L.M., Hellstrand, E., White, D.A., Rajah, L., Otzen, D.E., Vendruscolo, M., Dobson, C.M. and Knowles, T.P. (2013). Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. Proceedings of the National Academy of Sciences *110*, 9758–9763.

86. Rhoades, E., Agarwal, J. and Gafni, A. (2000). Aggregation of an amyloidogenic fragment of human islet amyloid polypeptide. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology *1476*, 230–238.

87. Rhoades, E. and Gafni, A. (2003). Micelle Formation by a Fragment of Human Islet Amyloid Polypeptide. Biophysical Journal *84*, 3480–3487.

88. Soreghan, B., Kosmoski, J. and Glabe, C. (1994). Surfactant properties of Alzheimer's A beta peptides and the mechanism of amyloid aggregation. The Journal of Biological Chemistry *18*, 28551–28554.

89. Powers, E.T. and Powers, D.L. (2008). Mechanisms of protein fibril formation: nucleated polymerization with competing off-pathway aggregation. Biophysical Journal *94*, 379–391.

90. Kamgar-Parsi, K., Hong, L., Naito, A., Brooks, C.L. and Ramamoorthy, A. (2017). Growth-incompetent monomers of human calcitonin lead to a noncanonical direct relationship between peptide concentration and aggregation lag time. Journal of Biological Chemistry *292*, 14963–14976.

91. Bothe, J.R., Andrews, A., Smith, K.J., Joyce, L.A., Krishnamachari, Y. and Kashi, S. (2019). Peptide Oligomerization Memory Effects and Their Impact on the Physical Stability of the GLP-1 Agonist Liraglutide. Molecular pharmaceutics *16*, 2153–2161.

92. Zapadka, K.L., Becher, F.J., Santos, A.L.G.D. and Jackson, S.E. (2017). Factors affecting the physical stability (aggregation) of peptide therapeutics. Interface focus *7*, 20170030.

93. Kamerzell, T.J., Esfandiary, R., Joshi, S.B., Middaugh, C.R. and Volkin, D.B. (2011). Protein–excipient interactions: Mechanisms and biophysical characterization applied to protein formulation development. Advanced Drug Delivery Reviews *63*, 1118–1159.

94. Pikal-Cleland, K. and Rodríguez-Hornedo, N. (2000). Protein denaturation during freezing and thawing in phosphate buffer systems: monomeric and tetrameric β-galactosidase. Archives of Biochemistry and Biophysics *384*, 398–406.

95. Erak, M., Bellmann-Sickert, K., Els-Heindl, S. and Beck-Sickinger, A.G. (2018). Peptide chemistry toolbox - Transforming natural peptides into peptide therapeutics. Bioorganic & medicinal chemistry *26*, 2759–2765.

96. Baets, G.D., Schymkowitz, J. and Rousseau, F. (2014). Predicting aggregation-prone sequences in proteins. Essays in Biochemistry *56*, 41–52.

97. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. and Dobson, C.M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. Letters to Nature *424*, 805–808.

98. Bolognesi, B. and Tartaglia, G.G. (2013). Chapter Three - Physicochemical Principles of Protein Aggregation. In Progress in Molecular Biology and Translational Science, 53–72.

99. Agrawal, N.J., Kumar, S., Wang, X., Helk, B., Singh, S.K. and Trout, B.L. (2011). Aggregation in Protein-Based Biotherapeutics: Computational Studies and Tools to Identify Aggregation-Prone Regions. Journal of Pharmaceutical Sciences *100*, 5081–5095.

100. Pawar, A.P., DuBay, K.F., Zurdo, J., Chiti, F., Vendruscolo, M. and Dobson, C.M. (2005). Prediction of “Aggregation-prone” and “Aggregation-susceptible” Regions in Proteins Associated with Neurodegenerative Diseases. Journal of Molecular Biology *350*, 379–392.

101. DuBay, K.F., Pawar, A.P., Chiti, F., Zurdo, J., Dobson, C.M. and Vendruscolo, M. (2004). Prediction of the Absolute Aggregation Rates of Amyloidogenic Polypeptide Chains. Journal of Molecular Biology *341*, 1317–1326.

102. Zibaee, S., Makin, O.S., Goedert, M. and Serpell, L.C. (2007). A simple algorithm locates β-strands in the amyloid fibril core of α-synuclein, Aβ, and tau using the amino acid sequence alone. Protein Science *16*, 906–918.

103. Beerten, J., Jonckheere, W., Rudyak, S., Xu, J., Wilkinson, H., Smet, F.D., Schymkowitz, J. and Rousseau, F. (2012). Aggregation gatekeepers modulate

protein homeostasis of aggregating sequences and affect bacterial fitness. Protein Engineering Design and Selection *25*, 357–366.

104. Baets, G.D., Durme, J.V., Rousseau, F. and Schymkowitz, J. (2014). A Genome-Wide Sequence–Structure Analysis Suggests Aggregation Gatekeepers Constitute an Evolutionary Constrained Functional Class. Journal of Molecular Biology *426*, 2405–2412.

105. Reumers, J., Maurer-Stroh, S., Schymkowitz, J. and Rousseau, F. (2009). Protein sequences encode safeguards against aggregation. Human Mutation *30*, 431–437.

106. Tartaglia, G.G., Pawar, A.P., Campioni, S., Dobson, C.M., Chiti, F. and Vendruscolo, M. (2008). Prediction of Aggregation-Prone Regions in Structured Proteins. Journal of Molecular Biology *380*, 425–436.

107. Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. and Serrano, L. (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nature Biotechnology *22*, 1302–1306.

108. Galzitskaya, O.V., Garbuzynskiy, S.O. and Lobanov, M.Y. (2006). Prediction of Amyloidogenic and Disordered Regions in Protein Chains. PLoS Computational Biology *2*, e177.

109. Bui, J.M., Cavalli, A. and Gsponer, J. (2008). Identification of Aggregation-Prone Elements by Using Interaction-Energy Matrices. Angewandte Chemie International Edition *47*, 7267–7269.

110. Trovato, A., Seno, F. and Tosatto, S.C.E. (2007). The PASTA server for protein aggregation prediction. Protein Engineering Design and Selection *20*, 521–523.

111. Bryan, A.W., Menke, M., Cowen, L.J., Lindquist, S.L. and Berger, B. (2009). BETASCAN: Probable β-amyloids Identified by Pairwise Probabilistic Analysis. PLoS Computational Biology *5*, e1000333.

112. Thompson, M.J., Sievers, S.A., Karanicolas, J., Ivanova, M.I., Baker, D. and Eisenberg, D. (2006). The 3D profile method for identifying fibril-forming segments of proteins. Proceedings of the National Academy of Sciences *103*, 4074–4078.

113. Zhang, Z., Chen, H. and Lai, L. (2007). Identification of amyloid fibril-forming segments based on structure and residue-based statistical potential. Bioinformatics *23*, 2218–2225.

114. Maurer-Stroh, S., Debulpaep, M., Kuemmerer, N., Paz, M.L. de la, Martins, I.C., Reumers, J., Morris, K.L., Copland, A., Serpell, L., Serrano, L., Schymkowitz, J.W.H. and Rousseau, F. (2010). Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. Nature Methods *7*, 237–242.

115. Conchillo-Solé, O., Groot, N.S. de, Avilés, F.X., Vendrell, J., Daura, X. and Ventura, S. (2007). AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics *8*, 65.

116. Groot, N. de, Pallarés, I., Avilés, F.X., Vendrell, J. and Ventura, S. (2005). Prediction of "hot spots" of aggregation in disease-linked polypeptides. BMC Structural Biology *5*, 18.

117. Walsh, I., Seno, F., Tosatto, S.C.E. and Trovato, A. (2014). PASTA 2.0: an improved server for protein aggregation prediction. Nucleic Acids Research *42*, W301–W307.

118. Arosio, P., Knowles, T.P. and Linse, S. (2015). On the lag phase in amyloid fibril formation. Physical Chemistry Chemical Physics *17*, 7606–7618.

119. Meisl, G., Kirkegaard, J.B., Arosio, P., Michaels, T.C., Vendruscolo, M., Dobson, C.M., Linse, S. and Knowles, T.P. (2016). Molecular mechanisms of protein aggregation from global fitting of kinetic models. Nature Protocols *11*, 252–272.

120. Buell, A.K., Dobson, C.M. and Knowles, T.P.J. (2014). The physical chemistry of the amyloid phenomenon: thermodynamics and kinetics of filamentous protein aggregation. Essays in Biochemistry *56*, 11–39.

121. Gillam, J.E. and MacPhee, C.E. (2013). Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth. Journal of Physics: Condensed Matter *25*, 373101.

122. Hall, D., Kardos, J., Edskes, H., Carver, J.A. and Goto, Y. (2015). A multi-pathway perspective on protein aggregation: Implications for control of the rate and extent of amyloid formation. FEBS Letters *589*, 672–679.

123. Borzova, V.A., Markossian, K.A. and Kurganov, B.I. (2014). Relationship between the initial rate of protein aggregation and the lag period for amorphous aggregation. International Journal of Biological Macromolecules *68*, 144–150.

124. Borgia, M.B., Nickson, A.A., Clarke, J. and Hounslow, M.J. (2013). A Mechanistic Model for Amorphous Protein Aggregation of Immunoglobulin-like Domains. Journal of the American Chemical Society *135*, 6456–6464.

125. Morriss-Andrews, A. and Shea, J.-E. (2012). Kinetic pathways to peptide aggregation on surfaces: The effects of β-sheet propensity and surface attraction. The Journal of Chemical Physics *136*, 065103.

126. Brorsson, A.-C., Kumita, J.R., MacLeod, I., Bolognesi, B., Speretta, E., Luheshi, L.M., Knowles, T.P.J., Dobson, C.M. and Crowther, D.C. (2010). Methods and models in neurodegenerative and systemic protein aggregation diseases. Frontiers in bioscience (Landmark edition) *15*, 373–396.

127. Meisl, G., Yang, X., Frohm, B., Knowles, T.P. and Linse, S. (2016). Quantitative analysis of intrinsic and extrinsic factors in the aggregation mechanism of Alzheimer-associated Aβ-peptide. Scientific Reports *6*, 18728.

128. Iljina, M., Garcia, G.A., Horrocks, M.H., Tosatto, L., Choi, M.L., Ganzinger, K.A., Abramov, A.Y., Gandhi, S., Wood, N.W., Cremades, N., Dobson, C.M., Knowles, T.P.J. and Klenerman, D. (2016). Kinetic model of the aggregation of alpha-synuclein provides insights into prion-like spreading. Proceedings of the National Academy of Sciences *113*, E1206–E1215.

129. Shammas, S.L., Garcia, G.A., Kumar, S., Kjaergaard, M., Horrocks, M.H., Shivji, N., Mandelkow, E., Knowles, T.P.J., Mandelkow, E. and Klenerman, D. (2015). A mechanistic model of tau amyloid aggregation based on direct observation of oligomers. Nature Communications *6*, 7025.

130. Deva, T., Lorenzen, N., Vad, B.S., Petersen, S.V., Thørgersen, I., Enghild, J.J., Kristensen, T. and Otzen, D.E. (2013). Off-pathway aggregation can inhibit fibrillation at high protein concentrations. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1834*, 677–687.

131. Marshall, K.E., Morris, K.L., Charlton, D., O'Reilly, N., Lewis, L., Walden, H. and Serpell, L.C. (2011). Hydrophobic, aromatic, and electrostatic interactions play a central role in amyloid fibril formation and stability. Biochemistry *50*, 2061–2071.

132. Kamihira, M., Oshiro, Y., Tuzi, S., Nosaka, A.Y., Saitô, H. and Naito, A. (2003). Effect of Electrostatic Interaction on Fibril Formation of Human Calcitonin as Studied by High Resolution Solid State 13C NMR. Journal of Biological Chemistry *278*, 2859–2865.

133. Yun, S., Urbanc, B., Cruz, L., Bitan, G., Teplow, D.B. and Stanley, H.E. (2007). Role of Electrostatic Interactions in Amyloid β-Protein (Aβ) Oligomer Formation: A Discrete Molecular Dynamics Study. Biophysical Journal *92*, 4064–4077.

134. Raman, B., Chatani, E., Kihara, M., Ban, T., Sakai, M., Hasegawa, K., Naiki, H., Rao, C.M. hM and Goto, Y. (2005). Critical balance of electrostatic and hydrophobic interactions is required for beta2-microglobulin amyloid fibril growth and stability. Biochemistry *44*, 1288–1299.

135. Jain, S. and Udgaonkar, J.B. (2010). Salt-induced modulation of the pathway of amyloid fibril formation by the mouse prion protein. Biochemistry *49*, 7615–7624.

136. Klement, K., Wieligmann, K., Meinhardt, J., Hortschansky, P., Richter, W. and Fändrich, M. (2007). Effect of Different Salt Ions on the Propensity of Aggregation and on the Structure of Alzheimer's Aβ(1-40) Amyloid Fibrils. Journal of Molecular Biology *373*, 1321–1333.

137. Marek, P., Patsalo, V., Green, D. and Raleigh, D.P. (2012). Ionic strength effects on amyloid formation by amylin are a complicated interplay among

Debye screening, ion selectivity, and Hofmeister effects. Biochemistry *51*, 8478–8490.

138. Calamai, M., Kumita, J.R., Mifsud, J., Parrini, C., Ramazzotti, M., Ramponi, G., Taddei, N., Chiti, F. and Dobson, C.M. (2006). Nature and Significance of the Interactions between Amyloid Fibrils and Biological Polyelectrolytes. Biochemistry *45*, 12806–12815.

139. Zapadka, K.L. (2016). Understanding the mechanism of aggregation and amyloid fibril formation of glucagon-like peptide-1 (PhD Thesis, University of Cambridge).

140. Torosantucci, R., Schöneich, C. and Jiskoot, W. (2013). Oxidation of Therapeutic Proteins and Peptides: Structural and Biological Consequences. Pharmaceutical Research *31*, 541–553.

141. Wakankar, A.A. and Borchardt, R.T. (2006). Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. Journal of Pharmaceutical Sciences *95*, 2321–2336.

142. Ji, J.A., Zhang, B., Cheng, W. and Wang, Y.J. (2009). Methionine, tryptophan, and histidine oxidation in a model protein, PTH: Mechanisms and stabilization. Journal of Pharmaceutical Sciences *98*, 4485–4500.

143. Wang, W. (1999). Instability, stabilization, and formulation of liquid protein pharmaceuticals. International Journal of Pharmaceutics *185*, 129–188.

144. Nygren, H. and Stenberg, M. (1989). Immunochemistry at interfaces. Immunology *66*, 321–327.

145. Stenberg, M. and Nygren, H. (1988). Kinetics of antigen-antibody reactions at solid-liquid interfaces. Journal of Immunological Methods *113*, 3–15.

146. Wang, W., Nema, S. and Teagarden, D. (2010). Protein aggregation—Pathways and influencing factors. International Journal of Pharmaceutics *390*, 89–99.

147. Rabe, M., Verdes, D. and Seeger, S. (2011). Understanding protein adsorption phenomena at solid surfaces. Advances in Colloid and Interface Science *162*, 87–106.

148. Nygren, H. (1995). Logarithmic growth in surface adsorption. Advances in Colloid and Interface Science *62*, 137–159.

149. Nault, L., Vendrely, C., Bréchet, Y., Bruckert, F. and Weidenhaupt, M. (2013). Peptides that form β-sheets on hydrophobic surfaces accelerate surface-induced insulin amyloidal aggregation. FEBS Letters *587*, 1281–1286.

150. Nault, L., Guo, P., Jain, B., Bréchet, Y., Bruckert, F. and Weidenhaupt, M. (2013). Human insulin adsorption kinetics, conformational changes and

amyloidal aggregate formation on hydrophobic surfaces. Acta Biomaterialia *9*, 5070–5079.

151. Hamley, I.W. (2010). Interface influence. Nature Chemistry *2*, 707–708.

152. Sluzky, V., Tamada, J. and Klibanov, A. (1991). Kinetics of insulin aggregation in aqueous solutions upon agitation in the presence of hydrophobic surfaces. Proceedings of the National Academy of Sciences of the United States of America *88*, 9377–9381.

153. Pronchik, J., He, X., Giurleo, J.T. and Talaga, D.S. (2010). In Vitro Formation of Amyloid from α-Synuclein Is Dominated by Reactions at Hydrophobic Interfaces. Journal of the American Chemical Society *132*, 9797–9803.

154. Huotari, A., Xu, W., Mönkäre, J., Kovalainen, M., Herzig, K.-H., Lehto, V.-P. and Järvinen, K. (2013). Effect of surface chemistry of porous silicon microparticles on glucagon-like peptide-1 (GLP-1) loading, release and biological activity. International Journal of Pharmaceutics *454*, 67–73.

155. Mauri, S., Volk, M., Byard, S., Berchtold, H. and Arnolds, H. (2015). Stabilization of Insulin by Adsorption on a Hydrophobic Silane Self-Assembled Monolayer. Langmuir *31*, 8892–8900.

156. Pandey, L.M., Denmat, S.L., Delabouglise, D., Bruckert, F., Pattanayek, S.K. and Weidenhaupt, M. (2012). Surface chemistry at the nanometer scale influences insulin aggregation. Colloids and Surfaces B: Biointerfaces *100*, 69–76.

157. Feingold, V., Jenkins, A.B. and Kraegen, E.W. (1984). Effect of contact material on vibration-induced insulin aggregation. Diabetologia *27*, 373–378.

158. Sefton, M.V. and Antonacci, G.M. (1984). Adsorption Isotherms of Insulin onto Various Materials. Diabetes *33*, 674–680.

159. Lougheed, W.D., Woulfe-Flanagan, H., Clement, J.R. and Albisser, A.M. (1980). Insulin aggregation in artificial delivery systems. Diabetologia *19*, 1–9.

160. Moores, B., Drolle, E., Attwood, S.J., Simons, J. and Leonenko, Z. (2011). Effect of Surfaces on Amyloid Fibril Formation. PLoS ONE *6*, e25954.

161. Chi, E.Y., Frey, S.L., Winans, A., Lam, K.L.H., Kjaer, K., Majewski, J. and Lee, K.Y.C. (2010). Amyloid-β Fibrillogenesis Seeded by Interface-Induced Peptide Misfolding and Self-Assembly. Biophysical Journal *98*, 2299–2308.

162. Giacomelli, C.E. and Norde, W. (2005). Conformational Changes of the Amyloidβ-Peptide (1-40) Adsorbed on Solid Surfaces. Macromolecular Bioscience *5*, 401–407.

163. Giacomelli, C.E. and Norde, W. (2003). Influence of Hydrophobic Teflon Particles on the Structure of Amyloid β-Peptide. Biomacromolecules *4*, 1719–1726.

164. Kowalewski, T. and Holtzman, D.M. (1999). In situ atomic force microscopy study of Alzheimer's β-amyloid peptide on different substrates: New insights into mechanism of β-sheet formation. Proceedings of the National Academy of Sciences *96*, 3688–3693.

165. Pinholt, C., Kapp, S.J., Bukrinsky, J.T., Hostrup, S., Frokjaer, S., Norde, W. and Jorgensen, L. (2013). Influence of acylation on the adsorption of GLP-2 to hydrophobic surfaces. International Journal of Pharmaceutics *440*, 63–71.

166. Trier, S., Linderoth, L., Bjerregaard, S., Andresen, T.L. and Rahbek, U.L. (2014). Acylation of Glucagon-Like Peptide-2: Interaction with Lipid Membranes and In Vitro Intestinal Permeability. PLoS ONE *9*, e109939.

167. Lu, D.R. (1993). Glucagon adsorption on polymer surfaces with α-helical and extended β-strand conformations: A computational approach. Journal of Biomaterials Science, Polymer Edition *4*, 323–335.

168. Pinholt, C., Bukrinsky, J.T., Hostrup, S., Frokjaer, S., Norde, W. and Jorgensen, L. (2011). Influence of PEGylation with linear and branched PEG chains on the adsorption of glucagon to hydrophobic surfaces. European Journal of Pharmaceutics and Biopharmaceutics *77*, 139–147.

169. Rabe, M., Soragni, A., Reynolds, N.P., Verdes, D., Liverani, E., Riek, R. and Seeger, S. (2013). On-Surface Aggregation of α-Synuclein at Nanomolar Concentrations Results in Two Distinct Growth Mechanisms. ACS Chemical Neuroscience *4*, 408–417.

170. Schladitz, C., Vieira, E.P., Hermel, H. and Möhwald, H. (1999). Amyloid–β-Sheet Formation at the Air-Water Interface. Biophysical Journal *77*, 3305–3310.

171. Mauri, S., Weidner, T. and Arnolds, H. (2014). The structure of insulin at the air/water interface: monomers or dimers? Physical Chemistry Chemical Physics *16*, 26722–26724.

172. Hlady, V. and Buijs, J. (1996). Protein adsorption on solid surfaces. Current Opinion in Biotechnology *7*, 72–77.

173. Zbacnik, T.J., Holcomb, R.E., Katayama, D.S., Murphy, B.M., Payne, R.W., Coccaro, R.C., Evans, G.J., Matsuura, J.E., Henry, C.S. and Manning, M.C. (2017). Role of Buffers in Protein Formulations. Journal of pharmaceutical sciences *106*, 713–733.

174. Widmaier, E.P., Raff, H. and Strang, K.T. (2008). Vander's Human Physiology: The Mechanisms of Body Function 11th ed. (McGraw-Hill Higher Education).

175. Pedersen, J.S., Dikov, D. and Otzen, D.E. (2006). N- and C-Terminal Hydrophobic Patches Are Involved in Fibrillation of Glucagon. Biochemistry *45*, 14503–14512.

176. Abelein, A., Jarvet, J., Barth, A., Gräslund, A. and Danielsson, J. (2016). Ionic Strength Modulation of the Free Energy Landscape of Aβ40 Peptide Fibril Formation. Journal of the American Chemical Society *138*, 6893–6902.

177. Bee, J.S., Randolph, T.W., Carpenter, J.F., Bishop, S.M. and Dimitrova, M.N. (2011). Effects of Surfaces and Leachables on the Stability of Biopharmaceuticals. Journal of Pharmaceutical Sciences *100*, 4158–4170.

178. Khan, T.A., Mahler, H.-C. and Kishore, R.S.K. (2015). Key interactions of surfactants in therapeutic protein formulations: A review. European Journal of Pharmaceutics and Biopharmaceutics *97*, 60–67.

179. Lee, H.J., McAuley, A., Schilke, K.F. and McGuire, J. (2011). Molecular origins of surfactant-mediated stabilization of protein drugs. Advanced Drug Delivery Reviews *63*, 1160–1171.

180. Nidhi, K., Indrajeet, S., Khushboo, M., Gauri, K. and Sen, D.J. (2010). Hydrotropy: A Promising Tool for Solubility Enhancement: A Review. International Journal of Drug Development and Research *3*, 26–33.

181. Kishore, R.S.K., Kiese, S., Fischer, S., Pappenberger, A., Grauschopf, U. and Mahler, H.-C. (2011). The Degradation of Polysorbates 20 and 80 and its Potential Impact on the Stability of Biotherapeutics. Pharmaceutical Research *28*, 1194–1210.

182. Kishore, R.S.K., Pappenberger, A., Dauphin, I.B., Ross, A., Buergi, B., Staempfli, A. and Mahler, H.-C. (2011). Degradation of Polysorbates 20 and 80: Studies on Thermal Autoxidation and Hydrolysis. Journal of Pharmaceutical Sciences *100*, 721–731.

183. Tomlinson, A., Demeule, B., Lin, B. and Yadav, S. (2015). Polysorbate 20 Degradation in Biopharmaceutical Formulations: Quantification of Free Fatty Acids, Characterization of Particulates, and Insights into the Degradation Mechanism. Molecular Pharmaceutics *12*, 3805–3815.

184. Wang, W., Wang, Y.J. and Wang, D.Q. (2008). Dual effects of Tween 80 on protein stability. International Journal of Pharmaceutics *347*, 31–38.

185. Ha, E., Wang, W. and Wang, Y.J. (2002). Peroxide formation in polysorbate 80 and protein stability. Journal of Pharmaceutical Sciences *91*, 2252–2264.

186. Maggio, E.T. (2010). Use of excipients to control aggregation in peptide and protein formulations. Journal of Excipients and Food Chemicals *2*, 40–49.

187. Otzen, D. (2011). Protein–surfactant interactions: A tale of many states. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1814*, 562–591.

188. Macchi, F., Eisenkolb, M., Kiefer, H. and Otzen, D. (2012). The effect of osmolytes on protein fibrillation. Molecular Sciences *13*, 3801–3819.

189. Lam, X.M., Yang, J.Y. and Cleland, J.L. (1997). Antioxidants for Prevention of Methionine Oxidation in Recombinant Monoclonal Antibody HER2. Journal of Pharmaceutical Sciences *86*, 1250–1255.

190. Tsai, P.K., Volkin, D.B., Dabora, J.M., Thompson, K.C., Bruner, M.W., Gress, J.O., Matuszewska, B., Keogan, M., Bondi, J.V. and Middaugh, C.R. (1993). Formulation Design of Acidic Fibroblast Growth Factor. Pharmaceutical Research *10*, 649–659.

191. Zhou, S., Zhang, B., Sturm, E., Teagarden, D.L., Schöneich, C., Kolhe, P., Lewis, L.M., Muralidhara, B.K. and Singh, S.K. (2010). Comparative Evaluation of Disodium Edetate and Diethylenetriaminepentaacetic Acid as Iron Chelators to Prevent Metal-Catalyzed Destabilization of a Therapeutic Monoclonal Antibody. Journal of Pharmaceutical Sciences *99*, 4239–4250.

192. Meyer, B.K., Ni, A., Hu, B. and Shi, L. (2007). Antimicrobial preservative use in parenteral products: Past and present. Journal of Pharmaceutical Sciences *96*, 3155–3167.

193. Zhang, Y., Roy, S., Jones, L.S., Krishnan, S., Kerwin, B.A., Chang, B.S., Manning, M.C., Randolph, T.W. and Carpenter, J.F. (2004). Mechanism for benzyl alcohol-induced aggregation of recombinant human interleukin-1 receptor antagonist in aqueous solution. Journal of Pharmaceutical Sciences *93*, 3076–3089.

194. Singh, S.M., Hutchings, R.L. and Mallela, K.M.G. (2011). Mechanisms of m-cresol-induced Protein Aggregation Studied Using a Model Protein Cytochrome c. Journal of Pharmaceutical Sciences *100*, 1679–1689.

195. Teska, B.M., Alarcón, J., Pettis, R.J., Randolph, T.W. and Carpenter, J.F. (2014). Effects of Phenol and meta-Cresol Depletion on Insulin Analog Stability at Physiological Temperature. Journal of Pharmaceutical Sciences *103*, 2255–2267.

196. Bye, J.W., Platts, L. and Falconer, R.J. (2014). Biopharmaceutical liquid formulation: a review of the science of protein stability and solubility in aqueous environments. Biotechnology Letters *36*, 869–875.

197. Ohtake, S., Kita, Y. and Arakawa, T. (2011). Interactions of formulation excipients with proteins in solution and in the dried state. Advanced Drug Delivery Reviews *63*, 1053–1073.

198. Lee, J. (2000). Biopharmaceutical formulation. Current Opinion in Biotechnology *11*, 81–84.

199. Finder, V.H., Vodopivec, I., Nitsch, R.M. and Glockshuber, R. (2010). The Recombinant Amyloid-β Peptide Aβ1–42 Aggregates Faster and Is More Neurotoxic than Synthetic Aβ1–42. Journal of Molecular Biology *396*, 9–18.

200. Pedersen, J.S., Flink, J.M., Dikov, D. and Otzen, D.E. (2006). Sulfates Dramatically Stabilize a Salt-Dependent Type of Glucagon Fibrils. Biophysical Journal *90*, 4181–4194.

201. Cordeiro, Y., Foguel, D. and Silva, J.L. (2013). Pressure–temperature folding landscape in proteins involved in neurodegenerative diseases and cancer. Biophysical Chemistry *183*, 9–18.

202. Seefeldt, M., Rosendahl, M., Cleland, J. and Hesterberg, L. (2009). Application of High Hydrostatic Pressure to Dissociate Aggregates and Refold Proteins. Current Pharmaceutical Biotechnology *10*, 447–455.

203. Liu, R., Su, R., Liang, M., Huang, R., Wang, M., Qi, W. and He, Z. (2012). Physicochemical Strategies for Inhibition of Amyloid Fibril Formation: An Overview of Recent Advances. Current Medicinal Chemistry *19*, 4157–4174.

204. Wang, W. and Roberts, C.J. (2013). Non-Arrhenius Protein Aggregation. The AAPS Journal *15*, 840–851.

205. Mishra, R. and Winter, R. (2008). Cold- and Pressure-Induced Dissociation of Protein Aggregates and Amyloid Fibrils. Angewandte Chemie International Edition *47*, 6518–6521.

206. Kim, H.-Y., Cho, M.-K., Riedel, D., Fernandez, C.O. and Zweckstetter, M. (2008). Dissociation of Amyloid Fibrils of α-Synuclein in Supercooled Water. Angewandte Chemie International Edition *47*, 5046–5048.

207. Tobitani, A. and Ross-Murphy, S.B. (1997). Heat-Induced Gelation of Globular Proteins. 2. Effect of Environmental Factors on Single-Component and Mixed-Protein Gels. Macromolecules *30*, 4855–4862.

208. Giehm, L., Lorenzen, N. and Otzen, D.E. (2011). Assays for α-synuclein aggregation. Methods *53*, 295–305.

209. Batzli, K.M. and Love, B.J. (2015). Agitation of amyloid proteins to speed aggregation measured by ThT fluorescence: A call for standardization. Materials Science and Engineering: C *48*, 359–364.

210. Cicerone, M.T., Pikal, M.J. and Qian, K.K. (2015). Stabilization of proteins in solid form. Advanced Drug Delivery Reviews *93*, 14–24.

211. Roy, I. and Gupta, M.N. (2004). Freeze-drying of proteins: some emerging concerns. Biotechnology and Applied Biochemistry *39*, 165–177.

212. Jain, K., Salamat-Miller, N. and Taylor, K. (2021). Freeze–thaw characterization process to minimize aggregation and enable drug product manufacturing of protein based therapeutics. Scientific Reports *11*, 11332.

213. Cao, E., Chen, Y., Cui, Z. and Foster, P.R. (2003). Effect of freezing and thawing rates on denaturation of proteins in aqueous solutions. Biotechnology and Bioengineering *82*, 684–690.

214. Arsiccio, A., Giorsello, P., Marenco, L. and Pisano, R. (2020). Considerations on Protein Stability During Freezing and Its Impact on the Freeze-Drying Cycle: A Design Space Approach. Journal of Pharmaceutical Sciences *109*, 464–475.

215. Qvit, N., Rubin, S.J.S., Urban, T.J., Mochly-Rosen, D. and Gross, E.R. (2017). Peptidomimetic therapeutics: scientific approaches and opportunities. Drug Discovery Today *22*, 454–462.

216. Furman, J.L., Chiu, M. and Hunter, M.J. (2014). Early Engineering Approaches to Improve Peptide Developability and Manufacturability. The AAPS Journal *17*, 111–120.

217. Hagihara, Y. and Saerens, D. (2014). Engineering disulfide bonds within an antibody. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1844*, 2016–2023.

218. Pace, C.N., Scholtz, J.M. and Grimsley, G.R. (2014). Forces stabilizing proteins. FEBS Letters *588*, 2177–2184.

219. Vinther, T.N., Kjeldsen, T.B., Jensen, K.J. and Hubálek, F. (2015). The road to the first, fully active and more stable human insulin variant with an additional disulfide bond. Journal of Peptide Science *21*, 797–806.

220. Li, Y., Zheng, X., Tang, L., Xu, W. and Gong, M. (2011). GLP-1 analogs containing disulfide bond exhibited prolonged half-life in vivo than GLP-1. Peptides *32*, 1303–1312.

221. Avadisian, M. and Gunning, P.T. (2013). Extolling the benefits of molecular therapeutic lipidation. Molecular BioSystems *9*, 2179–2188.

222. Aicart-Ramos, C., Valero, R.A. and Rodriguez-Crespo, I. (2011). Protein palmitoylation and subcellular trafficking. Biochimica et biophysica acta *1808*, 2981–2994.

223. Li, Y., Wang, Y., Wei, Q., Zheng, X., Tang, L., Kong, D. and Gong, M. (2015). Variant fatty acid-like molecules Conjugation, novel approaches for extending the stability of therapeutic peptides. Scientific reports *5*, 18039.

224. Ward, B.P., Ottaway, N.L., Perez-Tilve, D., Ma, D., Gelfanov, V.M., Tschöp, M.H. and DiMarchi, R.D. (2013). Peptide lipidation stabilizes structure to enhance biological function. Molecular Metabolism *2*, 468–479.

225. Meece, J. (2009). Pharmacokinetics and Pharmacodynamics of Liraglutide, a Long-Acting, Potent Glucagon-Like Peptide-1 Analog. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy *29*, 33S-42S.

226. Gilroy, C.A., Luginbuhl, K.M. and Chilkoti, A. (2016). Controlled release of biologics for the treatment of type 2 diabetes. Journal of Controlled Release *240*, 151–164.

227. Wang, Y., Lomakin, A., Kanai, S., Alex, R. and Benedek, G.B. (2015). Transformation of Oligomers of Lipidated Peptide Induced by Change in pH. Molecular Pharmaceutics *12*, 411–419.

228. Vad, B., Thomsen, L.A., Bertelsen, K., Franzmann, M., Pedersen, J.M., Nielsen, S.B., Vosegaard, T., Valnickova, Z., Skrydstrup, T., Enghild, J.J., Wimmer, R., Nielsen, N.C. and Otzen, D.E. (2010). Divorcing folding from function: How acylation affects the membrane-perturbing properties of an antimicrobial peptide. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1804*, 806–820.

229. Privé, G.G. (2009). Lipopeptide detergents for membrane protein studies. Current Opinion in Structural Biology *19*, 379–385.

230. Gunn, A.P., Masters, C.L. and Cherny, R.A. (2010). Pyroglutamate-Aβ: Role in the natural history of Alzheimer's disease. The International Journal of Biochemistry & Cell Biology *42*, 1915–1918.

231. Goldblatt, G., Matos, J.O., Gornto, J. and Tatulian, S.A. (2015). Isotope-edited FTIR reveals distinct aggregation and structural behaviors of unmodified and pyroglutamylated amyloid β peptides. Physical Chemistry Chemical Physics *17*, 32149–32160.

232. Thal, D.R., Walter, J., Saido, T.C. and Fändrich, M. (2014). Neuropathology and biochemistry of Aβ and its aggregates in Alzheimer's disease. Acta Neuropathologica *129*, 167–182.

233. Rezaei-Ghaleh, N., Amininasab, M., Kumar, S., Walter, J. and Zweckstetter, M. (2016). Phosphorylation modifies the molecular stability of β-amyloid deposits. Nature Communications *7*, 11359.

234. Osaki, D. and Hiramatsu, H. (2016). Citrullination and deamidation affect aggregation properties of amyloid β-proteins. Amyloid *23*, 234–241.

235. Guivernau, B., Bonet, J., Valls-Comamala, V., Bosch-Morató, M., Godoy, J.A., Inestrosa, N.C., Perálvarez-Marín, A., Fernández-Busquets, X., Andreu, D., Oliva, B. and Muñoz, F.J. (2016). Amyloid-β Peptide Nitrotyrosination Stabilizes Oligomers and Enhances NMDAR-Mediated Toxicity. The Journal of Neuroscience *36*, 11693–11703.

236. Johnson, E.C.B., Lanning, J.D. and Meredith, S.C. (2016). Peptide backbone modification in the bend region of amyloid-β inhibits fibrillogenesis but not oligomer formation. Journal of Peptide Science *22*, 368–373.

237. van Diggelen, F., Tepper, A.W.J.W., Apetri, M.M. and Otzen, D.E. (2017). α-Synuclein Oligomers: A Study in Diversity. Israel Journal of Chemistry *57*, 699–723.

238. Broersen, K., Brink, D. van den, Fraser, G., Goedert, M. and Davletov, B. (2006). α-Synuclein Adopts an α-Helical Conformation in the Presence of Polyunsaturated Fatty Acids To Hinder Micelle Formation. Biochemistry *45*, 15610–15616.

239. Deleersnijder, A., Gerard, M., Debyser, Z. and Baekelandt, V. (2013). The remarkable conformational plasticity of alpha-synuclein: blessing or curse? Trends in Molecular Medicine *19*, 368–377.

240. Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., … Nussbaum, R.L. (1997). Mutation in the α-Synuclein Gene Identified in Families with Parkinson's Disease. Science *276*, 2045–2047.

241. Krüger, R., Kuhn, W., Müller, T., Woitalla, D., Graeber, M., Kösel, S., Przuntek, H., Epplen, J.T., Schols, L. and Riess, O. (1998). AlaSOPro mutation in the gene encoding α-synuclein in Parkinson's disease. Nature Genetics *18*, 106–108.

242. Zarranz, J.J., Alegre, J., Gómez-Esteban, J.C., Lezcano, E., Ros, R., Ampuero, I., Vidal, L., Hoenicka, J., Rodriguez, O., Atarés, B., Llorens, V., Tortosa, E.G., Ser, T. del, Muñoz, D.G. and Yebenes, J.G. de (2004). The new mutation, E46K, of α-synuclein causes parkinson and Lewy body dementia. Annals of Neurology *55*, 164–173.

243. Bartels, T., Kim, N.C., Luth, E.S. and Selkoe, D.J. (2014). N-Alpha-Acetylation of α-Synuclein Increases Its Helical Folding Propensity, GM1 Binding Specificity and Resistance to Aggregation. PLoS ONE *9*, e103727.

244. Iyer, A., Roeters, S.J., Schilderink, N., Hommersom, B., Heeren, R.M.A., Woutersen, S., Claessens, M.M.A.E. and Subramaniam, V. (2016). The Impact of N-terminal Acetylation of α-Synuclein on Phospholipid Membrane Binding and Fibril Structure. Journal of Biological Chemistry *291*, 21110–21122.

245. Xiang, W., Schlachetzki, J.C.M., Helling, S., Bussmann, J.C., Berlinghof, M., Schäffer, T.E., Marcus, K., Winkler, J., Klucken, J. and Becker, C.-M. (2013). Oxidative stress-induced posttranslational modifications of alpha-synuclein: Specific modification of alpha-synuclein by 4-hydroxy-2-nonenal increases dopaminergic toxicity. Molecular and Cellular Neuroscience *54*, 71–83.

246. Bae, E.-J., Ho, D.-H., Park, E., Jung, J.W., Cho, K., Hong, J.H., Lee, H.-J., Kim, K.P. and Lee, S.-J. (2013). Lipid Peroxidation Product 4-Hydroxy-2-Nonenal Promotes Seeding-Capable Oligomer Formation and Cell-to-Cell Transfer of α-Synuclein. Antioxidants & Redox Signaling *18*, 770–783.

247. Carmo-Gonçalves, P., Pinheiro, A.S., Romão, L., Cortines, J. and Follmer, C. (2014). UV-induced selective oxidation of Met5 to Met-sulfoxide leads to the formation of neurotoxic fibril-incompetent α-synuclein oligomers. Amyloid *21*, 163–174.

248. Chen, F., David, D., Ferrari, A. and Gotz, J. (2004). Posttranslational Modifications of Tau - Role in Human Tauopathies and Modeling in Transgenic Animals. Current Drug Targets *5*, 503–515.

249. Lin, C., Chen, E.H.-L., Lee, L.Y.-L., Hsu, R.-L., Luh, F.Y., Yang, L., Chou, C.-F., Huang, L.-D., Lin, C.-C. and Chen, R.P.-Y. (2014). Comparison of the anti-amyloidogenic effect of O-mannosylation, O-galactosylation, and O-GalNAc glycosylation. Carbohydrate Research *387*, 46–53.

250. Rawat, A. and Nagaraj, R. (2013). Covalently attached fatty acyl chains alter the aggregation behavior of an amyloidogenic peptide derived from human β2-microglobulin. Journal of Peptide Science *19*, 770–783.

251. Martin, B.R. and Cravatt, B.F. (2009). Large-scale profiling of protein palmitoylation in mammalian cells. Nature Methods *6*, 135–138.

252. Rajendran, L., Schneider, A., Schlechtingen, G., Weidlich, S., Ries, J., Braxmeier, T., Schwille, P., Schulz, J.B., Schroeder, C., Simons, M., Jennings, G., Knölker, H.-J. and Simons, K. (2008). Efficient inhibition of the Alzheimer's disease beta-secretase by membrane targeting. Science *320*, 520–3.

253. Johannessen, L., Remsberg, J., Gaponenko, V., Adams, K.M., Barchi, J.J., Tarasov, S.G., Jiang, S. and Tarasova, N.I. (2011). Peptide Structure Stabilization by Membrane Anchoring and its General Applicability to the Development of Potent Cell-Permeable Inhibitors. ChemBioChem *12*, 914–921.

254. Plum, A., Jensen, L.B. and Kristensen, J.B. (2013). In vitro protein binding of liraglutide in human plasma determined by reiterated stepwise equilibrium dialysis. Journal of Pharmaceutical Sciences *102*, 2882–2888.

255. Knudsen, L.B., Nielsen, P.F., Huusfeldt, P.O., Johansen, N.L., Madsen, K., Pedersen, F.Z., Thøgersen, H., Wilken, M. and Agersø, H. (2000). Potent Derivatives of Glucagon-like Peptide-1 with Pharmacokinetic Properties Suitable for Once Daily Administration. Journal of Medicinal Chemistry *43*, 1664–1669.

256. Day, J.W., Ottaway, N., Patterson, J.T., Gelfanov, V., Smiley, D., Gidda, J., Findeisen, H., Bruemmer, D., Drucker, D.J., Chaudhary, N., … Tschöp, M.H. (2009). A new glucagon and GLP-1 co-agonist eliminates obesity in rodents. Nature Chemical Biology *5*, 749–757.

257. Lund, P.K., Goodman, R.H., Dee, P.C. and Habener, J.F. (1982). Pancreatic preproglucagon cDNA contains two glucagon-related coding sequences arranged in tandem. Proceedings of the National Academy of Sciences *79*, 345–349.

258. Bell, G.I., Sanchez-Pescador, R., Laybourn, P.J. and Najarian, R.C. (1983). Exon duplication and divergence in the human preproglucagon gene. Nature *304*, 368–371.

259. Ørskov, C., Holst, J.J., Knuhtsen, S., Baldissera, S., Poulsen, F.G.A. and Nielsen, O.V. (1986). Glucagon-Like Peptides GLP-1 and GLP-2, Predicted Products of the Glucagon Gene, Are Secreted Separately from Pig Small Intestine but Not Pancreas. Endocrinology *119*, 1467–1475.

260. Drucker, D.J., Habener, J.F. and Holst, J.J. (2017). Discovery, characterization, and clinical development of the glucagon-like peptides. Journal of Clinical Investigation *127*, 4217–4227.

261. Ørskov, C., Bersani, M., Johnsen, A., Højrup, P. and Holst, J.J. (1989). Complete sequences of glucagon-like peptide-1 from human and pig small intestine. The journal of biological chemistry *264*, 12826–12829.

262. Holst, J. (2019). From the Incretin Concept and the Discovery of GLP-1 to Today's Diabetes Therapy. Frontiers in Endocrinology *10*, 260.

263. Kreymann, B., Ghatei, M.A., Williams, G. and Bloom, S.R. (1987). Glucagon-like peptide-1 7-36: a physiological incretin in man. The Lancet *330*, 1300–1304.

264. Ørskov, C., Holst, J.J. and Nielsen, O.V. (1988). Effect of Truncated Glucagon-Like Peptide-1 [Proglucagon-(78–107) amide] on Endocrine Secretion from Pig Pancreas, Antrum, and Nonantral Stomach. Endocrinology *123*, 2009–2013.

265. Holst, J.J. (1999). Glucagon-like Peptide 1 (GLP-1): An Intestinal Hormone, Signalling Nutritional Abundance, with an Unusual Therapeutic Potential. Trends in Endocrinology & Metabolism *10*, 229–235.

266. Drucker, D.J. and Nauck, M.A. (2006). The incretin system: glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors in type 2 diabetes. Lancet *368*, 1696–1705.

267. Deacon, C.F., Johnsen, A.H. and Holst, J.J. (1995). Degradation of glucagon-like peptide-1 by human plasma in vitro yields an N-terminally truncated peptide that is a major endogenous metabolite in vivo. The Journal of Clinical Endocrinology & Metabolism *80*, 952–957.

268. Ahren, B., Gomis, R., Standl, E., Mills, D. and Schweizer, A. (2004). Twelve- and 52-Week Efficacy of the Dipeptidyl Peptidase IV Inhibitor LAF237 in Metformin-Treated Patients With Type 2 Diabetes. Diabetes Care *27*, 2874–2880.

269. Deacon, C.F., Knudsen, L.B., Madsen, K., Wiberg, F.C., Jacobsen, O. and Holst, J.J. (1998). Dipeptidyl peptidase IV resistant analogues of glucagon-like peptide-1 which have extended metabolic stability and improved biological activity. Diabetologia *41*, 271–278.

270. Deacon, C.F., Pridal, L., Klarskov, L., Olesen, M. and Holst, J.J. (1996). Glucagon-like peptide 1 undergoes differential tissue-specific metabolism in

the anesthetized pig. American Journal of Physiology-Endocrinology and Metabolism *271*, E458–E464.

271. Eng, J., Kleinman, W.A., Singh, L., Singh, G. and Raufman, J.-P. (1992). Isolation and Characterization of Exendin-4, an Exendin-3 Analogue, from Heloderma suspecturn Venom. The Journal of Biological Chemistry *267*, 7402–7405.

272. Kolterman, O.G., Kim, D.D., Shen, L., Ruggles, J.A., Nielsen, L.L., Fineman, M.S. and Baron, A.D. (2005). Pharmacokinetics, pharmacodynamics, and safety of exenatide in patients with type 2 diabetes mellitus. American Journal of Health-System Pharmacy *62*, 173–181.

273. Schmidt, L.J., Habacher, W., Augustin, T., Krahulec, E. and Semlitsch, T. (2014). A systematic review and meta-analysis of the efficacy of lixisenatide in the treatment of patients with type 2 diabetes. Diabetes, Obesity and Metabolism *16*, 769–779.

274. Adelhorst, K., Hedegaard, B. and Knudsen, L. (1994). Structure-activity studies of glucagon-like peptide-1. The Journal of Biological Chemistry *269*, 6275–6278.

275. Pi-Sunyer, X., Astrup, A., Fujioka, K., Greenway, F., Halpern, A., Krempf, M., Lau, D.C.W., Roux, C.W. le, Ortiz, R.V., Jensen, C.B., Wilding, J.P.H. and Group, S.O. and P.N.-1839 S. (2015). A Randomized, Controlled Trial of 3.0 mg of Liraglutide in Weight Management. New England Journal of Medicine *373*, 11–22.

276. Bush, M., Matthews, J. and Boever, D.E. (2009). Safety, tolerability, pharmacodynamics and pharmacokinetics of albiglutide, a long-acting glucagon-like peptide-1 mimetic, in healthy subjects. Diabetes, Obesity and Metabolism *11*, 498–505.

277. Matthews, J., Stewart, M., Boever, E.H.D., Dobbins, R.L., Rebecca, Walker, S.E., Holland, M.C., Bush, M.A. and Group, A.S. (2008). Pharmacodynamics, pharmacokinetics, safety, and tolerability of albiglutide, a long-acting glucagon-like peptide-1 mimetic, in patients with type 2 diabetes. The Journal of Clinical Endocrinology & Metabolism *93*, 4810–4817.

278. Lee, S. and Lee, D.Y. (2017). Glucagon-like peptide-1 and glucagon-like peptide-1 receptor agonists in the treatment of type 2 diabetes. Annals of Pediatric Endocrinology & Metabolism *22*, 15–26.

279. Lau, J., Bloch, P., Schäffer, L., Pettersson, I., Spetzler, J., Kofoed, J., Madsen, K., Knudsen, L.B., McGuire, J., Steensgaard, D.B., Strauss, H.M., Gram, D.X., Knudsen, S.M., Nielsen, F.S., Thygesen, P., Reedtz-Runge, S. and Kruse, T. (2015). Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide. Journal of Medicinal Chemistry *58*, 7370–7380.

280. Knowles, T.P., Waudby, C.A., Devlin, G.L., Cohen, S.I., Aguzzi, A., Vendruscolo, M., Terentjev, E.M., Welland, M.E. and Dobson, C.M. (2009). An Analytical Solution to the Kinetics of Breakable Filament Assembly. Science *326*, 1533–1537.

281. Vassar, P.S. and Culling, C.F. (1959). Fluorescent stains, with special reference to amyloid and connective tissues. Archives of pathology *68*, 487–498.

282. Naiki, H., Higuchi, K., Hosokawa, M. and Takeda, T. (1989). Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavine T. Analytical Biochemistry *177*, 244–249.

283. Biancalana, M. and Koide, S. (2010). Molecular mechanism of Thioflavin-T binding to amyloid fibrils. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1804*, 1405–1412.

284. Stsiapura, V.I., Maskevich, A.A., Kuzmitsky, V.A., Turoverov, K.K. and Kuznetsova, I.M. (2007). Computational Study of Thioflavin T Torsional Relaxation in the Excited State. The Journal of Physical Chemistry A *111*, 4829–4835.

285. Voropai, E.S., Samtsov, M.P., Kaplevskii, K.N., Maskevich, A.A., Stepuro, V.I., Povarova, O.I., Kuznetsova, I.M., Turoverov, K.K., Fink, A.L. and Uverskii, V.N. (2003). Spectral Properties of Thioflavin T and Its Complexes with Amyloid Fibrils. Journal of Applied Spectroscopy *70*, 868–874.

286. Sawaya, M.R., Sambashivan, S., Nelson, R., Ivanova, M.I., Sievers, S.A., Apostol, M.I., Thompson, M.J., Balbirnie, M., Wiltzius, J.J.W., McFarlane, H.T., Madsen, A.Ø., Riekel, C. and Eisenberg, D. (2007). Atomic structures of amyloid cross-β spines reveal varied steric zippers. Nature *447*, 453–457.

287. Nelson, R., Sawaya, M.R., Balbirnie, M., Madsen, A.Ø., Riekel, C., Grothe, R. and Eisenberg, D. (2005). Structure of the cross-β spine of amyloid-like fibrils. Nature *435*, 773–778.

288. Makin, O.S., Atkins, E., Sikorski, P., Johansson, J. and Serpell, L.C. (2005). Molecular basis for amyloid fibril formation and stability. Proceedings of the National Academy of Sciences of the United States of America *102*, 315–320.

289. Krebs, M.R.H., Bromley, E.H.C. and Donald, A.M. (2005). The binding of thioflavin-T to amyloid fibrils: localisation and implications. Journal of Structural Biology *149*, 30–37.

290. Wu, C., Wang, Z., Lei, H., Duan, Y., Bowers, M.T. and Shea, J.-E. (2008). The Binding of Thioflavin T and Its Neutral Analog BTA-1 to Protofibrils of the Alzheimer's Disease Aβ16–22 Peptide Probed by Molecular Dynamics Simulations. Journal of Molecular Biology *384*, 718–729.

291. Arad, E., Green, H., Jelinek, R. and Rapaport, H. (2020). Revisiting thioflavin T (ThT) fluorescence as a marker of protein fibrillation – the prominent role of electrostatic interactions. Journal of Colloid and Interface Science *573*, 87–95.

292. Biancalana, M., Makabe, K., Koide, A. and Koide, S. (2009). Molecular Mechanism of Thioflavin-T Binding to the Surface of β-Rich Peptide Self-Assemblies. Journal of Molecular Biology *385*, 1052–1063.

293. Wu, C., Biancalana, M., Koide, S. and Shea, J.-E. (2009). Binding Modes of Thioflavin-T to the Single-Layer β-Sheet of the Peptide Self-Assembly Mimics. Journal of Molecular Biology *394*, 627–633.

294. Chothia, C. (1973). Conformation of twisted β-pleated sheets in proteins. Journal of Molecular Biology *75*, 295–302.

295. Richardson, J.S. and Richardson, D.C. (2002). Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. Proceedings of the National Academy of Sciences *99*, 2754–2759.

296. Harel, M., Sonoda, L.K., Silman, I., Sussman, J.L. and Rosenberry, T.L. (2008). Crystal Structure of Thioflavin T Bound to the Peripheral Site of Torpedo californica Acetylcholinesterase Reveals How Thioflavin T Acts as a Sensitive Fluorescent Reporter of Ligand Binding to the Acylation Site. Journal of the American Chemical Society *130*, 7856–7861.

297. Sen, P., Fatima, S., Ahmad, B. and Khan, R.H. (2009). Interactions of thioflavin T with serum albumins: Spectroscopic analyses. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy *74*, 94–99.

298. Malmos, K.G., Blancas-Mejia, L.M., Weber, B., Buchner, J., Ramirez-Alvarado, M., Naiki, H. and Otzen, D. (2017). ThT 101: a primer on the use of thioflavin T to investigate amyloid formation. Amyloid *24*, 1–16.

299. Fauerbach, J.A., Yushchenko, D.A., Shahmoradian, S.H., Chiu, W., Jovin, T.M. and Jares-Erijman, E.A. (2012). Supramolecular Non-Amyloid Intermediates in the Early Stages of α-Synuclein Aggregation. Biophysical Journal *102*, 1127–1136.

300. Hackl, E.V., Darkwah, J., Smith, G. and Ermolina, I. (2015). Effect of acidic and basic pH on Thioflavin T absorbance and fluorescence. European Biophysics Journal *44*, 249–261.

301. Foderà, V., Groenning, M., Vetri, V., Librizzi, F., Spagnolo, S., Cornett, C., Olsen, L., Weert, M. van de and Leone, M. (2008). Thioflavin T Hydroxylation at Basic pH and Its Effect on Amyloid Fibril Detection. The Journal of Physical Chemistry B *112*, 15174–15181.

302. Singh, P.K., Mora, A.K. and Nath, S. (2015). Ultrafast fluorescence spectroscopy reveals a dominant weakly-emissive population of fibril bound thioflavin-T. Chemical Communications *51*, 14042–14045.

303. Coelho-Cerqueira, E., Pinheiro, A.S. and Follmer, C. (2014). Pitfalls associated with the use of Thioflavin-T to monitor anti-fibrillogenic activity. Bioorganic & Medicinal Chemistry Letters *24*, 3194–3198.

304. Wall, J., Murphy, C.L. and Solomon, A. (1999). In vitro immunoglobulin light chain fibrillogenesis. Methods in Enzymology *309*, 204–217.

305. Nafie, L.A. (1995). Circular polarization spectroscopy of chiral molecules. Journal of Molecular Structure *347*, 83–100.

306. Pain, R. (2004). Determining the CD Spectrum of a Protein. Current Protocols in Protein Science, 7.6.1-7.6.24.

307. Kelly, S.M., Jess, T.J. and Price, N.C. (2005). How to study proteins by circular dichroism. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1751*, 119–139.

308. Martin, S.R. and Schilstra, M.J. (2008). Circular Dichroism and Its Application to the Study of Biomolecules. Methods in Cell Biology *84*, 263–293.

309. Whitmore, L. and Wallace, B.A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. Biopolymers *89*, 392–400.

310. Kong, J. and Yu, S. (2007). Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. Acta Biochimica et Biophysica Sinica *39*, 549–559.

311. Jones, C., Mulloy, B., Thomas, A.H., Haris, P.I. and Chapman, D. (1993). Microscopy, Optical Spectroscopy, and Macroscopic Techniques. Methods in molecular biology *22*, 183–202.

312. Barth, A. (2007). Infrared spectroscopy of proteins. Biochimica et Biophysica Acta (BBA) - Bioenergetics *1767*, 1073–1101.

313. DeFlores, L.P., Ganim, Z., Nicodemus, R.A. and Tokmakoff, A. (2009). Amide I′−II′ 2D IR Spectroscopy Provides Enhanced Protein Secondary Structural Sensitivity. Journal of the American Chemical Society *131*, 3385–3391.

314. Kuttler, K. (2012). Computational Toxicology, Volume II. Methods in Molecular Biology *930*, 429–473.

315. Abdi, H. and Williams, L.J. (2010). Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics *2*, 433–459.

316. Zhang, L. and Wang, Y. (2014). Visualizing singular value decomposition. Wiley Interdisciplinary Reviews: Computational Statistics *6*, 197–201.

317. Graf, M. and Wätzig, H. (2004). Capillary isoelectric focusing – reproducibility and protein adsorption. Electrophoresis *25*, 2959–2964.

318. Loughney, J.W., Minsker, K., Ha, S. and Rustandi, R.R. (2019). Development of an imaged capillary isoelectric focusing method for characterizing the surface charge of mRNA lipid nanoparticle vaccines. Electrophoresis *40*, 2602–2609.

319. Rozing, G. (2019). Imaged Capillary Iso-Electric Focussing: Background, Status and Perspectives. Chromatography Today *February / March*, 8–14.

320. Kristl, T., Stutz, H., Wenz, C. and Rozing, G. (2014). Principles and Applications of Capillary Isoelectric Focusing.

321. Beynon, R. (2021). BioMol Buffer Calculator. https://www.biomol.net/en/tools/buffercalculator.htm.

322. Bioinformatics, S.I. of (2021). Expasy. Protein Parameter Computation. https://web.expasy.org/protparam/.

323. Sulatskaya, A.I., Lavysh, A.V., Maskevich, A.A., Kuznetsova, I.M. and Turoverov, K.K. (2017). Thioflavin T fluoresces as excimer in highly concentrated aqueous solutions and as monomer being incorporated in amyloid fibrils. Scientific Reports *7*, 2146.

324. Nielsen, L., Khurana, R., Coats, A., Frokjaer, S., Brange, J., Vyas, S., Uversky, V.N. and Fink, A.L. (2001). Effect of Environmental Factors on the Kinetics of Insulin Fibril Formation: Elucidation of the Molecular Mechanism. Biochemistry *40*, 6036–6046.

325. Alvarez-Martinez, M.-T., Fontes, P., Zomosa-Signoret, V., Arnaud, J.-D., Hingant, E., Pujo-Menjouet, L. and Liautard, J.-P. (2011). Dynamics of polymerization shed light on the mechanisms that lead to multiple amyloid structures of the prion protein. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1814*, 1305–1317.

326. Oosawa, F. and Asakura, S. (1975). Thermodynamics of the Polymerization of Protein (Academic Press).

327. Sreerama, N. and Woody, R.W. (1993). A self-consistent method for the analysis of protein secondary structure from circular dichroism. Analytical Biochemistry *209*, 32–44.

328. Sreerama, N. and Woody, R.W. (2004). Computation and analysis of protein circular dichroism spectra. Methods in Enzymology *383*, 318–351.

329. Sreerama, N. and Woody, R.W. (2000). Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. Analytical Biochemistry *287*, 252–260.

330. Sreerama, N., Venyaminov, S.Y.U. and Woody, R.W. (1999). Estimation of the number of α-helical and β-strand segments in proteins using circular dichroism spectroscopy. Protein Science *8*, 370–380.

331. Compton, L.A. and Johnson, W.C. (1986). Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. Analytical Biochemistry *155*, 155–167.

332. Manavalan, P. and Johnson, W.C. (1987). Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. Analytical Biochemistry *167*, 76–85.

333. Provencher, S.W. and Gloeckner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. Biochemistry *20*, 33–37.

334. Stokkum, I.H.M. van, Spoelder, H.J.W., Bloemendal, M., Grondelle, R. van and Groen, F.C.A. (1990). Estimation of protein secondary structure and error analysis from circular dichroism spectra. Analytical Biochemistry *191*, 110–118.

335. Cardamone, M. and Puri, N. (1992). Spectrofluorimetric assessment of the surface hydrophobicity of proteins. Biochemical Journal *282*, 589–593.

336. Nečas, D. and Physics, K.-P. (2012). Gwyddion: an open-source software for SPM data analysis. Central European Journal of Physics *10*, 181–188.

337. Moreno-Herrero, F. and Gomez-Herrero, J. (2012). Atomic Force Microscopy in Liquid: Biological Applications - Part 1 AFM: Basic Concepts. In Atomic Force Microscopy in Liquid: Biological Applications, First Edition, A. M. Baró and R. G. Reifenberger, eds., 3–34.

338. Bock, R.K. and Krischer, W. (1998). The Data Analysis BriefBook. Accelerator Physics.

339. Kozlowski, L.P. (2016). IPC–isoelectric point calculator. Biology Direct *11:55*.

340. Lide, D.R. (1991). Handbook of Chemistry and Physics (CRC Press).

341. Brichtová, E. (2020). Studies of aggregation and physical stability of GLP-1 analogues (First Year Report, University of Cambridge).

342. Fasman (1975). Section C Lipids Carbohydrates & Steroids. In Handbook of Biochemistry Volume I. (CRC Press).

343. Nick, M., Wu, Y., Schmidt, N., Prusiner, S., Stöhr, J. and DeGrado, W.F. (2018). A long-lived Aβ oligomer resistant to fibrillization. Biopolymers *109*, e23096.

344. Henry, S., Vignaud, H., Bobo, C., Decossas, M., Lambert, O., Harte, E., Alves, I.D., Cullin, C. and Lecomte, S. (2014). Interaction of Aβ1–42

Amyloids with Lipids Promotes “Off-Pathway” Oligomerization and Membrane Damage. Biomacromolecules *16*, 944–950.

345. Combet, S., Cousin, F., Rezaei, H. and Acta, N.-S. et (2019). Membrane interaction of off-pathway prion oligomers and lipid-induced on-pathway intermediates during prion conversion: A clue for neurotoxicity. BBA - Biomembranes *1861*, 514–523.

346. Kjaergaard, M., Dear, A.J., Kundel, F., Qamar, S., Meisl, G., Knowles, T.P.J.P. and Klenerman, D. (2018). Oligomer Diversity during the Aggregation of the Repeat Region of Tau. ACS Chemical Neuroscience *9*, 3060–3071.

347. Hasecke, F., Miti, T., Perez, C., Barton, J., Schölzel, D., Gremer, L., Grüning, C.S.R., Matthews, G., Meisl, G., Knowles, T.P.J., Willbold, D., Neudecker, P., Heise, H., Ullah, G., Hoyer, W. and Muschol, M. (2018). Origin of metastable oligomers and their effects on amyloid fibril self-assembly. Chemical Science *9*, 5937–5948.

348. Crespo, R., Villar-Alvarez, E. and Taboada, P. (2016). What can the kinetics of amyloid fibril formation tell about off-pathway aggregation? The Journal of Biological Chemistry *291*, 2018–2032.

349. Lee, M.-C.C., Yu, W.-C.C., Shih, Y.-H.H., Chen, C.-Y.Y., Guo, Z.-H.H., Huang, S.-J.J., Chan, J.C.C.C. and Chen, Y.-R.R. (2018). Zinc ion rapidly induces toxic, off-pathway amyloid-β oligomers distinct from amyloid-β derived diffusible ligands in Alzheimer’s disease. Scientific reports *8*, 4772.

350. Chufán, E.E., De, M., Eipper, B.A., Mains, R.E. and Amzel, M.L. (2009). Amidation of Bioactive Peptides: The Structure of the Lyase Domain of the Amidating Enzyme. Structure *17*, 965–973.

351. Zapadka, K.L., Becher, F.J., Santos, A.L.G.D. and Jackson, S.E. Characterisation of the conditions under which off-pathway oligomers form during the aggregation of the therapeutic peptide GLP-1.

352. Grigolato, F. and Arosio, P. (2019). Sensitivity analysis of the variability of amyloid aggregation profiles. Physical Chemistry Chemical Physics *21*, 1435–1442.

353. Cohen, S.I.A., Vendruscolo, M., Dobson, C.M. and Knowles, T.P.J. (2011). Nucleated polymerization with secondary pathways. II. Determination of self-consistent solutions to growth processes described by non-linear master equations. The Journal of Chemical Physics *135*, 065106.

354. Cohen, S.I.A., Vendruscolo, M., Dobson, C.M. and Knowles, T.P.J. (2011). Nucleated polymerization with secondary pathways. III. Equilibrium behavior and oligomer populations. *135*, 065107.

355. Vivian, J.T. and Callis, P.R. (2001). Mechanisms of Tryptophan Fluorescence Shifts in Proteins. Biophysical Journal *80*, 2093–2109.

356. Lindgren, M., Sörgjerd, K. and Hammarström, P. (2005). Detection and Characterization of Aggregates, Prefibrillar Amyloidogenic Oligomers, and Protofibrils Using Fluorescence Spectroscopy. Biophysical Journal *88*, 4200–4212.

357. Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas', A.F. and Gilmanshin, R.I. (1991). Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. Biopolymers *31*, 119–128.

358. Spence, G.R., Capaldi, A.P. and Radford, S.E. (2004). Trapping the On-pathway Folding Intermediate of Im7 at Equilibrium. Journal of Molecular Biology *341*, 215–226.

359. Hennessey, J.P. and Johnson, C.W. (1981). Information content in the circular dichroism of proteins. Biochemistry *20*, 1085–1094.

360. Novo, M., Freire, S. and Al-Soufi, W. (2018). Critical aggregation concentration for the formation of early Amyloid-β (1–42) oligomers. Scientific Reports *8*, 1783.

361. Matulis, D. and Lovrien, R. (1998). 1-Anilino-8-Naphthalene Sulfonate Anion-Protein Binding Depends Primarily on Ion Pair Formation. Biophysical Journal *74*, 422–429.

362. Malmos, K.G., Blancas-Mejia, L.M., Weber, B., Buchner, J., Ramirez-Alvarado, M., Naiki, H. and Otzen, D. (2017). ThT 101: a primer on the use of thioflavin T to investigate amyloid formation. Amyloid *24*, 1–16.

363. Frederiksen, T.M., Sønderby, P., Ryberg, L.A., Harris, P., Bukrinski, J.T., Scharff-Poulsen, A.M., Elf-Lind, M.N. and Peters, G.H.H. (2015). Oligomerization of a Glucagon-like Peptide 1 Analog: Bridging Experiment and Simulations. Biophysical journal *109*, 1202–13.

364. Agyei, D., Ahmed, I., Akram, Z., Iqbal, H.M.N. and Danquah, M.K. (2017). Protein and Peptide Biopharmaceuticals: An Overview. Protein & Peptide Letters *24*, 94–101.

365. Agyei, D., Tan, K.-X., Pan, S., Udenigwe, C.C. and Danquah, M.K. (2018). 9 - Peptides for biopharmaceutical applications. In Peptide Applications in Biomedicine, Biotechnology and Bioengineering, 231–251.

366. Hannappel, M. (2017). Biopharmaceuticals: From peptide to drug. *1871*, 060004.

367. Pan, C.Q., Buxton, J.M., Yung, S.L., Tom, I., Yang, L., Chen, H. and MacDougall, M. (2006). Design of a long acting peptide functioning as both a glucagon-like peptide-1 receptor agonist and a glucagon receptor antagonist. Journal of Biological Chemistry *281*, 12506–12515.

368. Frokjaer, S. and Otzen, D.E. (2005). Protein drug stability: a formulation challenge. Nature Reviews Drug Discovery *4*, 298–306.

369. Moorkens, E., Meuwissen, N., Huys, I., Declerck, P., Vulto, A.G. and Simoens, S. (2017). The Market of Biopharmaceutical Medicines: A Snapshot of a Diverse Industrial Landscape. Frontiers in pharmacology *8*, 314.

370. Assarsson, A., Hellstrand, E., Cabaleiro-Lago, C. and Linse, S. (2014). Charge Dependent Retardation of Amyloid β Aggregation by Hydrophilic Proteins. ACS Chemical Neuroscience *5*, 266–274.

371. Basava, C. and Hostetler, K. (1996). Lipid conjugates of therapeutic peptides and protease inhibitors. United States Patent, 5,554,728.

372. Sahin, E. and Roberts, C.J. (2012). Therapeutic Proteins, Methods and Protocols. Methods in Molecular Biology *899*, 403–423.

373. Bowers, M.T. (2014). Ion mobility spectrometry: A personal view of its development at UCSB. International Journal of Mass Spectrometry *370*, 75–95.

374. Ashcroft, A.E. (2010). Mass spectrometry and the amyloid problem—How far can we go in the gas phase? Journal of the American Society for Mass Spectrometry *21*, 1087–1096.

375. Gabelica, V., Shvartsburg, A.A., Afonso, C., Barran, P., Benesch, J.L.P., Bleiholder, C., Bowers, M.T., Bilbao, A., Bush, M.F., Campbell, J.L., … Wyttenbach, T. (2019). Recommendations for reporting ion mobility Mass Spectrometry measurements. Mass Spectrometry Reviews *38*, 291–320.

376. Lanucara, F., Holman, S.W., Gray, C.J. and Eyers, C.E. (2014). The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nature Chemistry *6*, 281–294.

377. Nettleton, E.J., Tito, P., Sunde, M., Bouchard, M., Dobson, C.M. and Robinson, C.V. (2000). Characterization of the Oligomeric States of Insulin in Self-Assembly and Amyloid Fibril Formation by Mass Spectrometry. Biophysical Journal *79*, 1053–1065.

378. Lord, R.S., Gubensek, F. and Rupley, J.A. (1973). Insulin self-association. Spectrum changes and thermodynamics. Biochemistry *12*, 4385–4392.

379. Bernstein, S.L., Dupuis, N.F., Lazo, N.D., Wyttenbach, T., Condron, M.M., Bitan, G., Teplow, D.B., Shea, J.-E., Ruotolo, B.T., Robinson, C.V. and Bowers, M.T. (2009). Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. Nature Chemistry *1*, 326–331.

380. Bernstein, S.L., Wyttenbach, T., Baumketner, A., Shea, J.-E., Bitan, G., Teplow, D.B. and Bowers, M.T. (2005). Amyloid β-Protein: Monomer

Structure and Early Aggregation States of Aβ42 and Its Pro19 Alloform. Journal of the American Chemical Society *127*, 2075–2084.

381. Jablonowska, A., Bakun, M., Kupniewska-Kozak, A. and Dadlez, M. (2004). Alzheimer's Disease Aβ Peptide Fragment 10–30 Forms a Spectrum of Metastable Oligomers with Marked Preference for N to N and C to C Monomer Termini Proximity. Journal of Molecular Biology *344*, 1037–1049.

382. Wang, Y., Lomakin, A., Kanai, S., Alex, R. and Benedek, G.B. (2017). Liquid–Liquid Phase Separation in Oligomeric Peptide Solutions. Langmuir *33*, 7715–7721.

383. Kim, J.-K., Anderson, J., Jun, H.-W., Repka, M.A. and Jo, S. (2009). Self-Assembling Peptide Amphiphile-Based Nanofiber Gel for Bioresponsive Cisplatin Delivery. Molecular Pharmaceutics *6*, 978–985.

384. Castelletto, V., Cheng, G., Stain, C., Connon, C.J. and Hamley, I.W. (2012). Self-Assembly of a Peptide Amphiphile Containing l-Carnosine and Its Mixtures with a Multilamellar Vesicle Forming Lipid. Langmuir *28*, 11599–11608.

385. Rexeisen, E.L., Fan, W., Pangburn, T.O., Taribagil, R.R., Bates, F.S., Lodge, T.P., Tsapatsis, M. and Kokkoli, E. (2010). Self-Assembly of Fibronectin Mimetic Peptide-Amphiphile Nanofibers. Langmuir *26*, 1953–1959.

386. Rodriguez, L.M.D.L., Hemar, Y., Cornish, J. and Brimble, M.A. (2016). Structure–mechanical property correlations of hydrogel forming β-sheet peptides. Chemical Society Reviews *45*, 4797–4824.

387. Banwell, E.F., Abelardo, E.S., Adams, D.J., Birchall, M.A., Corrigan, A., Donald, A.M., Kirkland, M., Serpell, L.C., Butler, M.F. and Woolfson, D.N. (2009). Rational design and application of responsive α-helical peptide hydrogels. Nature Materials *8*, 596–600.

388. Austerberry, J.I., Thistlethwaite, A., Fisher, K., Golovanov, A.P., Pluen, A., Esfandiary, R., Walle, C.F. van der, Warwicker, J., Derrick, J.P. and Curtis, R. (2019). Arginine to Lysine Mutations Increase the Aggregation Stability of a Single-Chain Variable Fragment through Unfolded-State Interactions. Biochemistry *58*, 3413–3421.

8 APPENDICES

ANALYTICAL DATA SHEETS

GLP-1

BACHEM
PIONEERING PARTNER FOR PEPTIDES

Analytical Data Sheet

Lot number **1061673**
not for drug use

Catalog number	**H-9560**
Product number	**4043014**
Product	**GLP- 1 (7- 37) (human, bovine, guinea pig, mouse, rat) acetate salt** H- His- Ala- Glu- Gly- Thr- Phe- Thr- Ser- Asp- Val- Ser- Ser- Tyr- Leu- Glu- Gly- Gln- Ala- Ala- Lys- Glu- Phe- Ile- Ala- Trp- Leu- Val- Lys- Gly- Arg- Gly- OH acetate salt
Molecular formula	$C_{151}H_{228}N_{40}O_{47}$
Relative molecular mass	3355.71

Tests	Results
Appearance	white lyophilisate
Appearance of solution	clear, colorless solution (1 mg/mL in 50% acetic acid)
Identification (ESI-MS)	m = 3355.7u (average mass)
Purity (HPLC)	99.2% (TFA) 96.6% (TEAP)
Assay (elemental analysis)	91.4% (Nth 16.70%, Nfd 15.27%)

Latest update: October 19, 2015

Analytical Data Sheet
Lot number **1061673**

This Analytical Data Sheet is automatically generated based on reviewed analytical data and is therefore unsigned. 1/1

Am-GLP-1

BACHEM
PIONEERING PARTNER FOR PEPTIDES

Analytical Data Sheet

Lot number **3014701**
not for drug use

Product number	**4094597**
Product	**H- His- Ala- Glu- Gly- Thr- Phe- Thr- Ser- Asp- Val- Ser- Ser- Tyr- Leu- Glu- Gly- Gln- Ala- Ala- Lys- Glu- Phe- Ile- Ala- Trp- Leu- Val- Lys- Gly- Arg- Gly- NH_2 acetate salt**
Molecular formula	$C_{151}H_{229}N_{41}O_{46}$
Relative molecular mass	3354.73

Tests	Results
Appearance	white powder
Identification (MALDI-MS)	3354.93
Amino acid analysis	Asx 1.04 (1) Ala 3.99 (4) Phe 1.87 (2) Thr 1.99 (2) Val 1.91 (2) His 1.03 (1) Ser 2.84 (3) Ile 0.97 (1) Lys 2.04 (2) Glx 4.16 (4) Leu 1.91 (2) Trp* 0.69 (1) Gly 4.07 (4) Tyr 0.99 (1) Arg 1.02 (1) *Partially destroyed during acid hydrolysis
Solubility	soluble in water at 1mg/mL
Purity (HPLC)	96.7%
Assay (AAA)	97.3%
Date of manufacture	February 19, 2016

Latest update: February 25, 2016

Analytical Data Sheet
Lot number **3014701**

This Analytical Data Sheet is automatically generated based on reviewed analytical data and is therefore unsigned. 1/1

IPP4

BACHEM
PIONEERING PARTNER FOR PEPTIDES

Analytical Data Sheet

Lot number **3014726**
not for drug use

Product number	**4094484**
Product	**H- His- Ala- Glu- Gly- Thr- Phe- Thr- Ser- Asp- Val- Ser- Ser- Tyr- Leu- Glu- Gly- Lys(γ- Glu- palmitoyl) - Ala- Ala- Arg- Glu- Phe- Ile- Ala- Trp- Leu- Val- Arg- Gly- Arg- Gly- NH_2 acetate salt**
Molecular formula	$C_{173}H_{270}N_{46}O_{49}$
Relative molecular mass	3778.33

Tests	Results
Appearance	white powder
Identification (MALDI-MS)	3778.35
Solubility	insoluble in water at 1mg/mL insoluble in DMF at 1mg/mL soluble in 50% Acetic acid at 1mg/mL
Purity (HPLC)	96.3%
Date of manufacture	March 04, 2016

Latest update: March 04, 2016

Analytical Data Sheet
Lot number **3014726**

This Analytical Data Sheet is automatically generated based on reviewed analytical data and is therefore unsigned. 1/1